National Library of Canada
Canadian Theses Service
Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Services des thèses canadiennes

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ MICROFILMÉE TELLE QUE NOUS L'AVONS REÇUE

NL-339 (r.86/06)

Canada

THE UNIVERSITY OF ALBERTA

ASPECTS OF IMAGE PROCESSING IN NOISY ENVIRONMENTS: RESTORATION, DETECTION, AND PATTERN RECOGNITION

by

Zhi-Qiang Liu

(C)

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF Doctor of Philosophy

DEPARTMENT OF ELECTRICAL ENGINEERING

EDMONTON, ALBERTA

SPRING 1987

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-37642-2

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR Zhi-Qiang Liu

TITLE OF THESIS ASPECTS OF IMAGE PROCESSING IN NOISY ENVIRONMENTS: RESTORATION, DETECTION, AND PATTERN RECOGNITION

DEGREE FOR WHICH THESIS WAS PRESENTED Doctor of Philosophy

YEAR THIS DEGREE GRANTED SPRING 1987

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) Zhi-Qiang Liu

PERMANENT ADDRESS:

Northwestern Polytechnical University, Xian, Shaanxi, P.R. China

DATED Dec. 5 1986

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled ASPECTS OF IMAGE PROCESSING IN NOISY ENVIRONMENTS: RESTORATION, DETECTION, AND PATTERN RECOGNITION submitted by Zhi-Qiang Liu in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

V. G. Gourishankar
.....
Supervisor

.....
.....
.....
.....
External Examiner

Date December 3 1986

To Bao-Sheng and Ru-Min

ABSTRACT

In one sense, the aim of computer vision is to extract and interpret the organization of images in ways which are useful in specific problem domains. However, such images taken in realistic environments inevitably suffer from various degradations including noise, blurring, geometrical distortion, etc.. These degradations pose many problems for implementing computer vision techniques. An extensive literature review has identified the interesting areas requiring further study. Of these areas, nonstationary image restoration, and image signal detection in noisy environments are pursued extensively in this thesis. Four processing algorithms and techniques have been developed in the framework of stochastical estimation theory.

In order to obtain perceptually more satisfying restored images, image fields are assumed to be nonstationary. Two restoration algorithms are derived, namely, a sequential Kalman filter and an adaptive iterative filter. The sequential filter is developed based on a causal state-space image model. The adaptive filter uses a modified received image model. Both algorithms include some local spatial activity measurements in the filter gains such that the restored images retain edge information. Simulations show that the visual quality of the restored images is significantly improved.

Matched filters have long been used for signal detection. They are optimal only when the noise is stationary and white. In order to effectively apply the matched filters to images embedded in nonstationary, nonwhite noise backgrounds, a new adaptive postfiltering technique is derived. Superiority of this technique is shown by experiments.

In the fourth algorithm a hierarchical approach for multi-object detection is presented. In this approach, the detection of object is divided into three steps: prefiltering, pattern recognition, and detection. The images are prefiltered to suppress noise which would otherwise affect the entire detecting operation. In pattern recognition, a linear least square mapping technique is applied for classification. In the detection part, a known object of a class is used to match the received image of the same class. It is shown by simulations that with this approach, the computation time is reduced by more than 50%.

Finally, suggestions for further reseach in the related areas are also included.

ACKNOWLEDGEMENTS

The author would like to express his deep gratitude to his supervisor, Professor V. G. Gourishankar, for his guidance in formulating the objectives of this work, and for his encouragement and support throughout the research. It was his wise counsel that brought the author to this research in the first place. His enormous patience and personal warmth have greatly inspired the author in his endeavours.

The author is heavily indebted to his mentor and friend, Killam Professor T. M. Caelli, for his help and guidance. In his capacity as a co-superviser, Professor Caelli has provided to the author very valuable advice which has had a significant impact on the major part of this dissertation. The professional excellence Professor Caelli inspires is unquestioned and is of enormous value, but his continued open-mindedness and genuine interest have been the most valuable gift of all.

The author wishes to thank the other two members of his Ph.D. supervisory committee, Professors Z. J. Koles and R. E. Rink, for their constructive criticism and many helpfull suggestions during the preparation of this thesis. The assistance the author received from Dr. Z. J. Koles, which facilitated the initial part of the research work, is greatly appreciated.

The author is very grateful to his parents, who gave him a desire for knowledge, and to his family whose patience and sacrifice have made this work possible.

The support and assistance the author received from the Department of Electrical Engineering and Dr. Caelli's Computer Vision Lab. are also gratefully acknowledged.

TABLE OF CONTENTS

CHAPTER		page
I.	GENERAL INTRODUCTION	1
	1.1 Perspective and Motivation	1
	1.2 Image Restoration	3
	1.3 Image Pattern Recognition and Detection	11
	1.4 Overview of the Research	16
II.	SEQUENTIAL KALMAN FILTER FOR IMAGE RESTORATION	20
	2.1 Introduction	20
	2.2 Image Modeling	22
	2.3 The Two-Dimensional Adaptive Kalman Filter	25
	2.4 The Local Spatial Statistics Measurements	34
	2.5 Identification of the Modeling Coefficients	39
	2.6 Simulations	42
III.	ADAPTIVE ITERATIVE IMAGE RESTORATION ALGORITHM	55
	3.1 Introduction	55
	3.2 Image Modeling	59
	3.3 The Restoration Algorithm	64
	3.4 Simulations	67
IV.	AN ADAPTIVE MATCHED FILTER FOR DETECTING IMAGES IN NONWHITE NOISE BACKGROUNDS	82
	4.1 Introduction	82
	4.2 Derivation of the Optimal Detection Filter	85
	4.3 The Signal Reference Adaptive Postfiltering Technique	91
	4.4 Detection Criteria: ROC Analysis	97
	4.5 Simulations	102
V.	A HIERARCHICAL APPROACH TO MULTI-OBJECT PATTERN RECOGNITION AND DETECTION IN NOISY BACKGROUNDS	109

5.1 Introduction 109

5.2 The Linear Least Squares Mapping Technique ..113

5.3 Estimation Scheme for the Prefilter 118

5.4 Implementation of the Hierarchical Detection Technique 122

5.5 Simulations 124

VI. CONCLUSIONS 146

6.1 Summary 146

6.2 Recommendations for Further Research 152

REFERENCES 156

APPENDIX I 163

APPENDIX II 169

APPENDIX III 171

LIST OF TABLES

Table		Page
3.1	Quantitative measurements of image quality	70
5.1	The mapping results for the training sample images	131
5.2	The mapping results for the noisy sample images	132
5.3	The mapping results for the noise-free object images	134
5.4	The mapping results for the noisy object images	135
5.5	The mapping results for the restored object images	136
5.6	The cross-correlation peak values	137

LIST OF FIGURES

Figure		Page
2.1	The Nonsymmetric Half Plane (NSHP)	23
2.2	The subregion of the NSHP for image modeling	23
2.3	The Square-Spiral scanning operation.	26
2.4	The notation of the causal field	26
2.5	The closest pixels for the state equation	26
2.6	An edge across the window region	37
2.7	The window divided into subregions R_{ij}	37
2.8	The local means for the subregions in the window and the orientation map	37
2.9	The Girl image	45
2.10	Test images and their histograms	47
2.11	The noisy and restored W.C. Fields images	49
2.12	The noisy and restored letter F images	51
2.13	The noisy and restored square pattern images	53
3.1	Noncausal image field	61
3.2	The window A_s used in the simulations	61
3.3	The technician image	71
3.4	The artificial pattern image	73
3.5	Cross section profiles of the images in Figure 3.4	75
3.6	Restored W.C. Fields images	76
3.7	Restored letter F images	78
3.8	Restored square pattern images	80
4.1	The reference adaptive postfiltering matching technique	94

Figure Page

4.2 Signal, its noise corrupted versions, the cross-correlations, and the ROC curves 98

4.3 The filter parameter configuration in a 5x5 window 103

4.4 Simulation results 107

5.1 The pattern recognition problem 112

5.2 Transformation of patterns into decision space 115

5.3 Patterns of class 1 centred around the point v_1 115

5.4 Mapping error caused by noise 115

5.5 The hierarchical detection approach 123

5.6 The flow diagram of the hierarchical detection approach 125

5.7 (a) the training sample images (three classes), (b) the noisy samples 138

5.8 (a) the object images, (b) the noisy objects (σ_v^2=1275) 140

5.9 (a) the restored object images, (b) the relative ROC curves for Bird 1 (object) 142

5.10 (a) the relative ROC curves for Face 2 (object), (b) the relative ROC curves for Leaf 3 (object) 144

CHAPTER I

GENERAL INTRODUCTION

1.1 Perspective and Motivation

Besides speech, images are a primary medium of human communication. A single image embodies an enormous quantity of information that cannot be carried, with the same concision, precision, and flexibility, by any other sensory medium. With the advent of the digital computer and new image processing techniques, we are able to obtain digitized images which are defined by finite arrays of intensity levels normally corresponding to natural visible information, and which use the digital images in a variety of ways for different purposes. The spatial information, when processed by computers, also provides us with an essential means for man-machine communication in automation and artificial intelligence. The present generation of high speed computers facilitates machine vision and image processing in an increasingly complex manner.

Techniques for digitizing images vary as much as the purposes of using such images. However, the common activities involved are basically the encoding, the storage, and the interpretation of information in the form of an image. The quality of an image is critical in many applications. For example, using video images as guiding media for industrial robots to detect and pick up particular

objects on conveyor belts may sometimes lead to erronous operation if the images are degraded (e.g. "snow flakes", blurring, and distortion).

X-rays play a very important role in clinic medicine and medical research. The quality of X-ray images is critical for detecting diseased area. However, due to the nonlinearities of imaging systems and low radiation energy, these types of images are often blurred and low in contrast, causing problems in diagnosis.

Detecting astronomical events is often accomplished through images observed by telescopes located on exploring satellites. However, because of transmission noise, air turbulance, disturbances caused by space radiation, and the finite aperture of the telescope, the observed images frequently suffer from various kinds of degradation which make further analysis more difficult.

Over the past three decades, such problems have motivated the development of many techniques for the restoration and noise reduction of images. One early fruitful application of image processing was at the Jet Propulsion Laboratory of the California Institute of Technology in early 1960's as part of the Apollo program. Since then, the development of better, more efficient and cheaper image encoders, analyzers, and display devices has made spatial information more readily available in many different digital forms for a variety of areas.

Although their applications are diverse, image restoration, detection, and pattern recognition follow the same basic mathematical principles. The theme of this dissertation is focused on the mathematical principles abstracted from a variety of problem areas. In the chapters to follow, three aspects of the problem shall be presented:

(1) Development of mathematical image models;

(2) Development of algorithms capable of restoring images or detecting objects in noisy environments given explicit criteria;

(3) Computer simulations to verify the proposed algorithms.

Some important techniques previously developed for image restoration and detection will be discussed briefly in the sequel (A representative sampling of the literature is given in the reference section.).

1.2 Image Restoration

Images acquired through recording media are often corrupted by various forms of degradation among which noise and blur are typical and the most prominent. The quality of a received image is limited by the receiver and the environment within which the image is taken. The image sensor noise is also a major factor. As the image sensor usually operates at its sensitivity limit, the noise associated will inevitably manifest itself in the received image, such as film grain noise and shot noise (Huang, 1966;

Barb, 1975; Andrews and Hunt, 1977). Blurring is often caused by the relative motion of objects and imaging systems, atmospheric turbulance, out of focus operation, physical limits of image sampling systems, etc..

Corrupted images often pose problems for analysis and detection. Techniques for improving visual quality of noisy images and recovering images from their degraded counterparts are referred to as image restoration and enhancement (Huang, 1981). Earlier research in this area was to develop filters in the frequency domain (Andrews, 1974; Andrews and Hunt, 1977; Pratt, 1978). Due to the limitations of these approaches the images were assumed to be linear and stationary processes, noise degradations were considered as broad band (white) or band limited, and the noise reduction techniques were carried out through the use of the discrete fast Fourier transform (FFT).

Extensions of one-dimensional Bayesian estimation technique to image restoration (Habibi, 1972) led to the development of stochastical restoration algorithms in the spatial domain. As will be seen, these techniques are, for the most part, variations of two-dimensional Kalman-type filters. In these algorithms the images are assumed to be random processes (usually stationary). The Kalman filter is a recursive estimation technique. In order to implement the Kalman filter in the spatial domain, we have to establish an artificial causal image model which often contains image

vectors of very large dimension, and due to the fact that any image is of finite extent we are unable to establish a global state space model as we do for the time-domain Kalman filter.

With a view to overcoming the high dimensionality and computational problems, Woods and his colleagues (1977, 1981, 1981a) proposed a special structure for image state vectors which takes only the closest pixels before and above the pixel being estimated. Based on these image vectors they were able to establish a local state-space representation required by the Kalman filter. To further reduce computational load they developed a so-called RUKF (reduced update Kalman filter) (Woods et al., 1976, 1979, 1981a) which is shown to be optimal in the sense that it minimizes the post update mean squares error, and only those pixels in the immediate vicinity are used for modeling. Similarly, Murphy and Silverman (1978) considered two other suboptimal restoration schemes, namely, strip restoration and constrained optimal restoration. In the latter scheme the gain matrix of the Kalman filter is constrained to be square.

The advantages of two-dimensional recursive filters are that they require less computation time than the nonrecursive ones, they can handle space-variant blur more easily, and, as is well known (Anderson and Moore, 1972) being Kalman-type filters they are capable of coping with

nonstationary problems.

Traditionally, for the ease of computation and mathematical tractability, an image is normally assumed to be a wide-sense stationary field. Therefore the statistical properties of the image are governed globally rather than locally by its stationary covariance matrix which is of Toeplitz form (Andrews and Hunt, 1977; Gray, 1971). Under this assumption the necessary computation can be carried out by FFT algorithms and consequently, the computation time is reduced drastically. The savings in computation time are, however, at the expense of restoration quality. Algorithms based on this assumption tend to globally smooth out the noise and any abrupt changes of image intensities as well. As a result, the restored images are smooth but blurred.

Recent studies in human vision (see for example, Caelli and Moraglia, 1986) have shown how edges are critical for pattern recognition and the perceptual registration of images. This observation suggests that the abrupt changes in image intensities should be retained as much as possible after the restoration in order to achieve good visual quality. It is intuitively obvious that the visual quality of the restored image can be improved if edges can be preserved. As discussed before, preserving edges while restoring the degraded image using stochastic filtering techniques is quite a contradictory task. Wallis (1976) suggested an algorithm based on local mean and variance. In

his approach each pixel is required to have the local average mean and variance, and sudden fluctuation is measured. The algorithm is able to preserve the edges in restoration. The main drawback of this technique is that it tends to over-enhance subtle details of the image at the expense of its principal features. Based on the modification of Wallis' model, Lee (1980,1981) developed adaptive algorithms with striking results. These techniques, though they lack the mathematical elegance and sophistication of a few other techniques, are simple and effective. Chan and Lim (1984) suggested an interesting and rather simple four-stage one-dimensional adaptive algorithm. In their approach they process an image in cascade along four directions (0°,45°,90°,135°). The basic idea is that if the one-dimensional filter is oriented in the same direction as the edge, the noise is filtered along the edge, and in doing so, the edges are preserved. The simulation results seem much better than those obtained from its two-dimensional counterpart. Other works on preserving edges can be found in the literature, e.g. Ingle, et al., 1976,1978; Nahi et al., 1975; Rajala, et al., 1981; Abramatic, et al., 1982, to mention a few.

Research on developing filters for nonstationary, and signal-dependent noise processes has recently been published by some authors. After investigating some statistical characteristics of image fields, Hunt et al. (1976; Trussell and Hunt, 1978) showed that images can be decomposed into

nonstationary mean Gaussian model and stationary fluctuations about the mean. The nonstationary mean is obtained by blurring any one of the ensemble image members, and it has the gross structure that represents the context of the ensemble. Lebedev and Mirkin (1975) proposed a composite image model that assumes that an image is composed of many different stationary components, and each has a distinct stationary correlation structure. Using this image model, Ingle and Woods (1976) applied the reduced update Kalmen filter (RUKF) to image restoration. They establish a bank of RUKF's running in parallel and use an identification-estimation scheme, in which each point is assigned a stationary image model and filtered by the specific Kalman filter. Most recently, Kuan et al. (1984) considered an image model with nonstationary mean and nonstationary variance (NMNV), and developed a recursive algorithm for image restoration. They leave the nonstationarity to the input process, so that the filter has a simple, space-invariant structure. In their scheme, it becomes unnecessary to construct a space-variant dynamic model for each image, and only the local mean and local variance of the original image need to be estimated.

It is known (Netravali, 1975; Anderson et al., 1976; Rajala, et al., 1981; Green, 1983) that the human visual system is insensitive to noise in high contrast regions of an image, and relatively more sensitive to noise in flat regions. Such perceptual criteria for the development of

image models and restoration algorithms can often be effective. For example, Anderson and Netravali (1976) defined a subjective error criterion based on the human visual system model and established an adaptive strategy that strikes a compromise between the loss of resolution and noise rejection such that the same amount of subjective noise is suppressed throughout the image. Abramatic and Silverman (1982) generalized this procedure in the framework of the classical Wiener filter. Motivated by the earlier work of others (Budrikis, 1972; Netravali et al., 1975; Anderson et al., 1976; Limb et al. 1978), Rajala and DeFigueiredo (1981) suggest a recursive least squares method for the restoration of an image distorted by blurs and corrupted by additive, white Gaussian noise. They combine a visibility function (Netravali, et al., 1975; Anderson, et al., 1976) with the Kalman filter previously developed by Aboutalib et al. (1972). They first segment the image into disjoint regions according to local spatial activity of the region, and then determine the covariance structures of these segments. In this framework, the filter adapts itself to each segment through the visibility function for nonstationary restoration.

For reasons of mathematical and computational convenience, the assumptions upon which the theoretical framework is established are often simplified and idealized. This sometimes leads to results that are far from being realistic. In the development of numerous image restoration

algorithms the noise is usually assumed to be additive, signal-independent, white, and Gaussian. For instance, the famous Wiener filter is facilitated by this stationarity assumption, while the most frequently used Kalman filter requires that the system be linear, finite-dimensional, and driven by white noise (Sage and Melsa, 1971). In contrast to these, many physical noise processes do not possess such nice properties. For example, a television vidicon sensor is a power-law device. That is, the output current is the illumination raised to a power. If this power is not unity then a nonlinear relation exists between the orginal signal and the measurement. Another cause of difficulty is the nonstationarity of many signals. For instance, the image recorded by a vidicon sensor is usually a highly structured entity and does not possess stationary statistics (Green, 1983). Further to this, the noise from a vidicon sensor is coupled into the signal by the shot-noise process of receiving individual photons. Such noise is signal-dependent and nonstationary. Many other physical noise processes such as Poisson noise and multiplicative noise are inherently signal-dependent. Blurring effect is in general space-variant. Consequently, the simplified assumptions which are often required by the existing filtering techniques may cause difficulty in practice.

Naderi and Sawchuk (1978) derived a nonstationary discrete Wiener filter for a signal-dependent film grain noise process. In addition to the generality of the noise

model, the filter is able to adapt itself to the local signal statistics given the conditional noise statistics. Recently, Kasturi et al.(1985) developed an adaptive point estimator for the same noise model. This filter is based on minimum mean squares error criterion, and the algorithm takes local mean to achieve the adaptive MMSE. Lo and Sawchuk (1979) derived a nonlinear MAP (maximum à postériori) filter for images degraded by Poisson noise. Ward and Saleh (1985) considered a particular and interesting situation where the signal is corrupted by additive noise and the random impulse response is a function of an independent noise process.

1.3 Image Pattern Recognition and Detection

Conventionally, robots are designed to implement pre-planned tasks within a fixed environment. To design robots with the ability to "see" and "decide" requires the use of pattern recognition techniques and decision theory. Many other areas such as medical image analysis, scene matching, target detection, etc. are also dependent upon the theories and techniques of pattern recognition and detection.

As a result of more than three decades of research, pattern recognition has become a distinct field (Niemann, 1979). In a narrow sense, image pattern recognition involves the classification of a given image pattern into one of a number of classes. This is normally done in two steps: a

given image is first transformed into a feature space, and then in the second step, classified. The objectives of establishing the feature space are twofold, namely, (a) reducing dimensionality and (b) establishing feature vectors. Since image vectors are, even for images of moderate sizes, very large, it is necessary to extract a smaller amount of data (which is called feature vector) to represent the original image. These extracted data carry the essential discriminatory information needed to solve the overall pattern recognition problem. In pattern classification, a recognition logic classifies the image patterns using the information embedded in the feature vectors (Sammon, 1970; Duda, et al., 1973)

Numerous papers exist in the literature on the feature extraction and most of them concentrate on finding orthogonal transformations (Ahmed and Rao, 1975) among which the Karhunen-Loéve or the eigenvector orthonormal expansion (Watanabe, 1965; Chien et al., 1967) is the most popular. It is well known that Karhunen-Loéve expansion is optimal with respect to data compression or fitting. Fukunaga and Koontz (1970) argued that Karhunen-Loéve expansion might not be optimal with respect to discriminating data. They suggested a modified Karhunen-Loéve expansion procedure in which a preliminary transformation for a two-class problem is carried out such that the eigenvectors which fit class one best are worst for representing class two. By selecting discriminatory feature samples, Foley and Sammon (1975)

established an optimal set of discriminant vectors for a two-class problem.

Another interesting area is pattern recognition by moment invariants. This idea was initially introduced by Hu (1962) who showed that the moment invariants calculated in terms of ordinary moments have the property that they are invariant when the image is shifted, scaled, or rotated. Since then considerable attention has been given to the study of moment invariants and their applications (Gonzalez, et al., 1977; Dudani, et al., 1977; Maitra, 1979; Teague, 1980). Using circular-harmonic expansion (Cormack, 1963), Hsu et al.(1982,1982a,1984) developed rotation-invariant pattern recognition methods. Recently, Abu-Mostafa and Psaltis (1985) generalized the moment invariant theory with a set of complex moments, with which they showed the relation between moment invariants and the circular harmonic expansion.

Although detection of signals has long been a topic in communication and information area, digital image detection is relatively new, and requires more attention as the visual media emerge. Image detection involves identifying the presence or the absence of known objects in a received image. As mentioned earlier, natural image signals inevitably suffer from various degradations, the most commonly encountered degradation being noise.

Detectors such as template matching (Pratt, 1978; Hall, 1979; Rosenfeld et al., 1982), and matched filters (Arcese, et al., 1970, Turin, 1976; Dessimoz, et al., 1984) are classical and fundamental developments in image detection. Over the years, these techniques have been used extensively for character recognition, target detection, map-matching, navigation etc..

Template matching compares a template to all possible search regions in image using some similarity measurements (Wong, et al., 1978; Munteanu, 1981) to locate the most likely position of the object. Obviously this kind of blind search is time consuming.

Several algorithms have been developed to alleviate the computational burden. Wong et al. (1978) and Munfeanu (1981) proposed some sequential methods to reduce the number of template pixels that must be matched to image pixels by accumulating the error between the template and search area pixel by pixel. Since the error accumulates rapidly, poor matches are detected quickly and rejected.

Using a hierarchical approach (Rosenfeld, et al., 1977; Tanimoto, 1982), the template matching is performed over the images of reduced resolution to locate possible positions of the objects. Although this approach simplifies the computation, this procedure becomes less sensitive and introduce more errors. Vanderbrug and Rosenfeld (1977)

introduced two-stage template matching in which they randomly select a subtemplate and perform matching the subtemplate in the first stage. In the second stage a full template is applied, but only at locations where a sufficient match between the subtemplate and the picture has been reached. With this coarse-fine approach, the matching procedure can be implemented more efficiently.

The object of both hierarchical and two-stage template matching is to minimize the number of locations at which the full template must be applied without missing the object or causing false alarms.

Although closely related to template matching, matched filters (or cross-correlators) are developed from a different view point. The original derivation can be found in many articles, in particular, the introductory paper by Turin (1976). It is assumed that a received noisy image $g(x,y)$ may contain a known object $f(x,y)$. The goal is to determine the presence of $f(x,y)$ by using certain detector $h(x,y)$. Designing the "best" detector is based on a well known SNR criterion. It has been shown (Pratt, 1978; Rosenfeld, et al., 1982) that for stationary, white noise processes the optimal detector $h(x,y)$ is, in fact, the known image itself. A filter of this form is referred to as the matched filter. In general, however, the form of h is more complicated and the computation is rather time consuming.

Because of their simplicity, matched filters are frequently used in practice even when the random processes are not stationary and white. They will often yield nonoptimal detection for a given SNR. Vijaya Kumar and Carroll (1984) recently studied the loss of optimality in using the matched filter when the noise is nonwhite. They derived the lower bound to measure the loss. Liu and Caelli (1986) developed an adaptive postfiltering technique which considerably improves the detectability of the matched filter in nonwhite noise environments. In this technique there is no need to compute the inverse of the large covariance matrix. However, the algorithm is suboptimal for a given SNR criterion.

1.4 Overview of the Research

This thesis is largely restricted to the discussion and development of computational techniques which may be applied to image restoration, detection, and pattern recognition. Computer vision and image processing have grown rapidly during last two decades, and to such an extent that, even with this restriction, the number of possible directions in which the work could be taken are many. However, since additive noise (with stationary or nonstationary statistics), as discussed in the last two sections, is one of the most common forms of degradations encountered in acquired images, our main attention, in this thesis, is focused upon processing images corrupted by additive noise.

In Chapter II, a state-space image model is established. Based on this model, a sequential recursive image restoration algorithm is developed. This algorithm is a Kalman-type filter which has been applied extensively to system estimation since the early 1960s. To preserve edge information in the image, a measure of pixel activity is used. Depending on the thresholds chosen, this algorithm can be adjusted to obtain restored images with better visual quality. Simulation results are included.

Chapter III presents the development of an iterative restoration algorithm. A new observed image model is proposed, in which the local statistical information is included. With this model and the local statistics measurement an adaptive image restoration algorithm is derived. Simulation results using the original observed image model and the modified one proposed here show that the algorithm using the latter model, as expected, preserves edge information and produces better visual quality. Because the logic behind edge preservation is that the noise in the vicinities of edges are less suppressed, the SNR's are slightly lower than those of non-adaptive counterparts.

In Chapter IV a signal reference adaptive matched filter is derived for the detection of known images embedded in nonwhite noise background. It is a common practice that the simple matched filters are used for detection of known objects, even if the background is nonwhite. This practice

leads to nonoptimal detection given the SNR criterion. As pointed out earlier in the introduction, using the exact optimal detector in nonwhite noise environments would be computationally impractical because of the high dimensionality of the covariance matrix, which has to be inverted. For some special cases when the random processes are stationary, the covariance matrices are of Toeplitz form. The inverse of such matrices may be relatively easily computed by FFT. However, in general, the inversion of large dimensional covariance matrices is necessary. The proposed algorithm preserves the simplicity of the matched filter while considerably improving the detectability.

In this algorithm, the form of a postfilter with unknown parameters is chosen. An identification algorithm is developed based on the least squares error between the signal auto-correlation and the postfiltered signal-to-image cross-correlation. Simulation results show that the proposed adaptive matched filter, though suboptimal, does significantly improve the detectability compared to the simple matched filter. The detectability is very close to that of the optimal filter, and inversion of covariance matrix is avoided. This detector can be applied to nonstationary, nonwhite noise environments.

Detecting multiple objects in a noisy environment is of particular importance to robot vision. Using template matching and matched filter techniques often involves

unnecessary searching procedures. It is more desirable if these filters could be applied only to a subscene in which an object considered is most likely present. In Chapter V, this problem is investigated. A hierarchical approach which uses both pattern recognition and matched filter is established.

An image scene containing multi-objects is fed into the the pattern recognizer where the object (objects) is classified to a proper category. Then the classified object (objects) of one category is cross-correlated with a known object of the same category to detect the existence of the particular object in the scene. A linear least squares mapping technique (LLSMT) is applied. The classifier is a logic that discriminates features by using some disciminant (or basis) functions. These functions are designed (or trained) by using a set of training samples which are ideal object images and are deterministic. In a noisy environment these pre-determined functions will often result in misclassification. In order to circumvent this problem a simple restoration algorithm is derived and used as the first stage to prefilter the noise. Simulation results are included.

Finally, a summary of the dissertation and suggestions for further investigation in these fields are included in Chapter VI.

CHAPTER II

SEQUENTIAL KALMAN FILTER FOR IMAGE RESTORATION

2.1 Introduction

Recently, recursive estimation techniques have been applied to image restoration in the hope of alleviating the high computational burden and obtaining optimal restoration.

Due to its high computational feasibility and efficiency, Kalman filter theory has long been applied to one-dimensional systems with great success. Therefore the extension of this application to two-dimensional systems is quite appealing. The research in this direction seems to have been first attempted by Nahi (1972) and Habibi (1972). Further work on development of two-dimensional Kalman filters has been carried out by many others, such as Powell and Silverman (1974), Woods and Radewan (1977), Murphy and Silverman (1978), Kondo et al.(1981), to mention a few. Some of the two-dimensional Kalman filters concern themselves with degradations due to noise (no blurring) only, while others attempt to handle both blurring and noise. Willsky (1978) has done a thorough survey of the related research. Woods and his colleagues (Woods and Radewan, 1977; Woods, 1979; Woods and Ingle, 1981; Kaufman, Woods, et al., 1983) have done the most extensive study yet of two-dimensional Kalman-type filter theory and techniques.

Conventionally, it is assumed that an image field is wide sense stationary (see APPENDIX I for detailed definitions). The statistical properties of the image field are globally characterized by its stationary covariance function. The structure of a stationary covariance matrix is of Toeplitz form which enables the use of FFT-based algorithms in restoration, significantly reducing the computation time. Unfortunately, in most cases, the image field does not possess stationary statistics, and restoration algorithms based on the stationarity assumption are insensitive to abrupt changes in the image intensity. Consequently, these restoration algorithms tend to smooth out the edges and the restored images lack details and are often blurred. It is plausible that images restored by using algorithms based on nonstationary statistics assumptions may have better visual quality. More recently, researchers have proposed different image models and procedures to handle the nonstationarity of the statistical characteristics of the images. Instead of assuming global stationarity for the image field, most of these algorithms assume that the image field is locally stationary. In their distinctive work on solving this problem, Hunt and his colleagues (1976; Trussell and Hunt, 1978) proposed a Gaussian probability density function with nonstationary mean for the image field. They also show that an image can be modeled as stationary fluctuations about a nonstationary ensemble mean which has the gross structure representing the context of

the ensemble. Further to this, Kuan et al. (1984) proposed the nonstationary mean and nonstationary variance (NMNV) image model.

In this chapter we shall consider a new adaptive Kalman-type filter for image restoration. This is developed based on point-wise scanning. The image is modeled by an autoregressive model. The sequential Kalman filter is developed in terms of the innovation process. In order to retain the edge information, a weighting function varying with the spatial activity is derived in the Kalman filter. Also used is a refined local spatial activity measurement, which enables the suppression of noise even in vicinities of edges.

2.2 Image Modeling

It is well known that for most images, significant correlation exists only between pixels close to each other (Rosenfeld and Kak, 1982). Therefore we shall model an image over a finite nonsymmetric half plane (NSHP) (Ekstrom and Woods, 1976) illustrated in Figure 2.1. For any spatial position of a scanner, all the previous processed pixels are considered as "past" states, while the pixel which is presently being processed is called the "present" state, and all the pixels to be processed are referred to as the "future" states. Obviously, this definition is rather artificial, as the scanning operation can take any order.

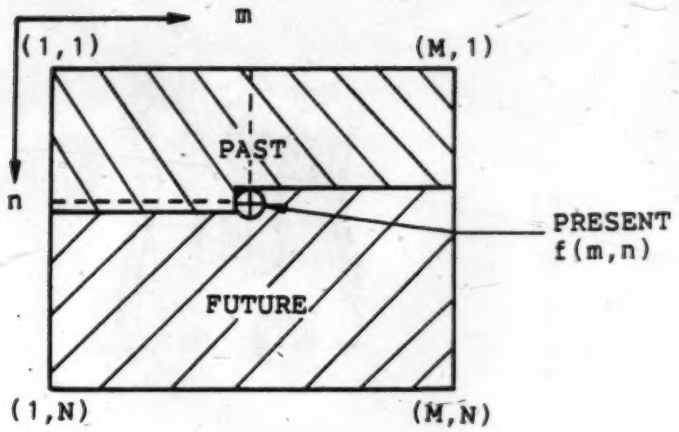


Figure 2.1 The Nonsymmetric Half Plane (NSHP)

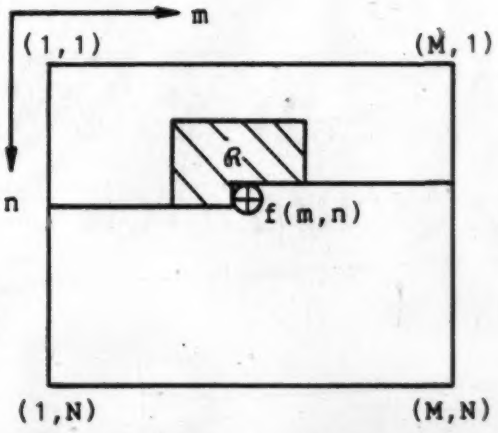


Figure 2.2 The subregion of the NSHP for image modeling

Let (m,n) be the present position of the scanner, we define $P(m,n)$ as past region with respect to (m,n),

$$P(m,n) = [(i,j):\{1\le i\le M,\ 1\le j\le n-1\}\cup\{1\le i\le m-1, j=n\}], \qquad (2.1)$$

where $\cup$ is union. The image signal at (m,n) can be modeled by an autoregressive model, namely,

$$f(m,n) = \sum_{(i,j)\in P(m,n)} a(i,j)f(m-i,n-j)+w_s(m,n), \qquad (2.2)$$

where w_s is modeling error and is assumed to be Gaussian and white,[1] and a(i,j) are the modeling coefficients.

Let g(m,n) represent the observed image which is corrupted by an additive noise v(m,n). We have

$$g(m,n) = f(m,n)+v(m,n), \qquad (2.3)$$

where v(m,n) is assumed to be a Gaussian, white noise array with zero mean and variance $\sigma_v^2(m,n)$, and is independent of f(m,n) and $w_s(m,n)$.

Usually the dimension of an image is very large, and so the simple use of Eq.(2.2) as the image model in Eq.(2.3) will lead to high dimensionality problem, and result in excessive computation. Motivated by the fact (Woods and

[1]This assumption is reasonable only for the autoregressive model, otherwise $w_s(m,n)$ will not be white (Nahi, 1972).

Radewan, 1977) that only those pixels closest to the present pixel are significant to the modeling of f(m,n), we reduce Eq.(2.2) to

$$f(m,n) = \sum_{(i,j)\in\mathcal{R}} a(i,j)f(m-i,n-j)+w_s(m,n), \tag{2.4}$$

where $\mathcal{R}\triangleq\{(k,l)\}$ is the sub-region of the NSHP shown in Figure 2.2, and w_s is the modeling error with zero mean and variance σ^2_{ws}. Without loss of generality, we assume $E[f(m,n)]=0$.

We have used a spiral scanning operation which runs continuously along a square-shaped spiral line (see Figure 2.3) and ends at the geometrical centre of the image.

Referring to Figure 2.4, for simplicity of notation we denote P as the "past", and F as the "future". For the particular spiral scanning configuration, the entire image can be considered as a long string of pixels. The image model (2.4) can be rewritten for the subregion Ω of P (see Figure 2.4) as follows,

$$f(m,n) = \sum_{(i,j)\in\Omega} a(i,j)f(m-i,n-j)+w_s(m,n). \tag{2.5}$$

2.3 The Two-Dimensional Adaptive Kalman Filter

The Kalman filtering technique is based on the linear minimum variance (or, sometimes, minimum mean squares error) estimate of f(m,n), given a linear observation g(m,n). This

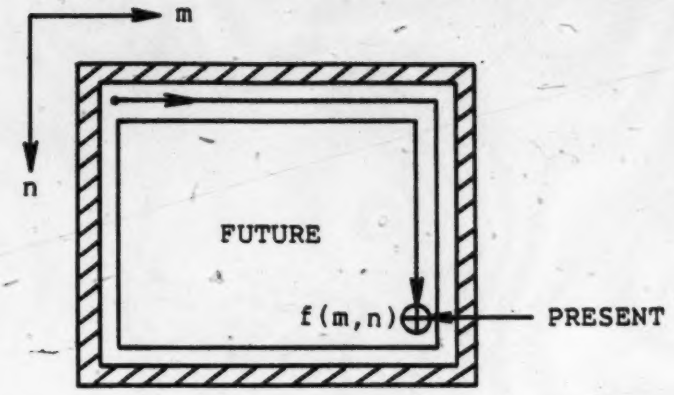


Figure 2.3 The Square-Spiral scanning operation

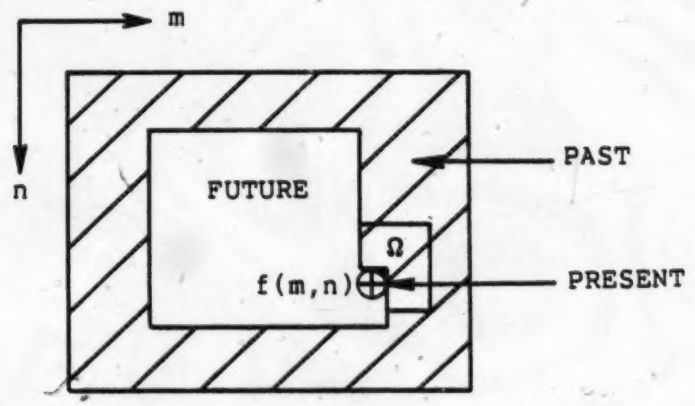


Figure 2.4 The notation of the causal field

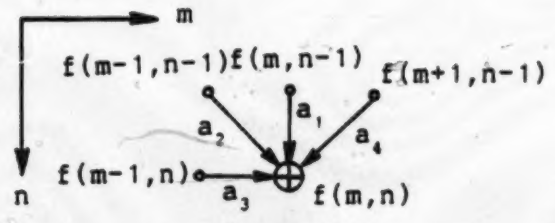


Figure 2.5 The closest pixels for the state equation

estimate is obtained by using the orthogonal projection principle (Kalman, 1960; Papoulis, 1984), which states that the linear minimum variance estimate $\hat{y}=Ax+B$ of y is such that the estimate error $(y-\hat{y})$ is orthogonal to the data x, or $\hat{y}$ is orthogonal projection of y onto X, in the sense that $E\{[y-\hat{y}]\}=0$, where $X=\{x\}$. In the present context, we require an estimate $\hat{f}$ of f given the observation g, such that $\hat{f}$ is the orthogonal projection of f onto P. That is, $\hat{f}=\hat{E}\{f|P\}$. Here we have used $\hat{E}$ rather than E to indicate that the linear minimum variance estimate is in general not the true conditional mean in the sense that the linear minimum variance estimate is a weighted linear combination of the observations.

The image models defined by Eqs.(2.5) and (2.3) over the causally partitioned image field provide the basis for all our succeeding development. In what follows, a problem of considerable importance shall be stated first.

Problem Statement

Given the models (2.3) and (2.5), determine an estimate $\hat{f}(m,n)$ (of f(m,n)) which is a linear combination of the observation g(m,n) and estimates $\{\hat{f}(i,j)\in\Omega\}$. The estimate is optimal subject to the criterion that the expected value of the sum of the variance between f and $\hat{f}$ is a minimum. That is, $\hat{f}(m,n)$ is to be chosen so that

$$\hat{E}\{[f(m,n)-\hat{f}(m,n)]^2\} = \text{minimum}. \qquad (2.6)$$

Signal Prediction and Update

We may write the estimate $\hat{f}(m,n)$ as

$$\hat{f}(m,n) = \hat{f}^{-}(m,n)+K(m,n)e(m,n), \tag{2.7}$$

where K(m,n) is an unknown weighting function (or the gain in the Kalman filter theory), the term $\hat{f}^{-}(m,n)$ defines the predicted estimate at (m,n), and is given by

$$\hat{f}^{-}(m,n) = \sum_{(i,j)\in\Omega} a(i,j)\hat{f}(m-i,n-j), \tag{2.8}$$

and

$$e(m,n) = g(m,n)-\hat{f}^{-}(m,n) \tag{2.9}$$

which is called the Kalman innovation process which plays an important role in the approaches to adaptive filtering techniques (Ljung and Söderström, 1983).

It is easy to show that, if the minimum variance estimator is unbiased (meaning that $E\{\hat{f}(m,n)\}=f(m,n)$), the innovation process is also white, with zero mean provided that the noise v is white with zero mean.

We may write Eq.(2.9) as follows

$$e(m,n) = \tilde{f}^{-}(m,n)+v(m,n), \tag{2.10}$$

where

$$\tilde{f}^{-}(m,n) = f(m,n)-\hat{f}^{-}(m,n).$$

By the orthogonality principle and from Eqs.(2.7) and (2.10) we have

$$E\{e(m,n)e(k,l)\} = E\{v(m,n)v(k,l)\} = 0, \ \forall m\neq k, \ n\neq l, \tag{2.11}$$

and obviously, since $E\{w_s(m,n)\} = 0$,

$$E\{e(m,n)\} = 0. \tag{2.12}$$

Equation (2.11) implies that e(.,.) is a white process, which may be written, in general, as

$$E\{e(m,n)e(k,l)\} = [V_f^{-}(m,n;m,n)+\sigma_v^2(m,n)]\delta_{mk}\delta_{nl} = R_e(m,n)\delta_{mk}\delta_{nl}, \tag{2.13}$$

where $R_e(m,n) \triangleq [V_f^{-}(m,n;m,n)+\sigma_v^2(m,n)]$ and

$$V_f^{-}(m,n;k,l) \triangleq E\{\tilde{f}^{-}(m,n)\tilde{f}^{-}(k,l)\} \tag{2.14}$$

which is the à priori error variance, and δ_{rs} is the Kronecker delta function,

$$\delta_{rs}=\begin{cases}1, & r=s,\\ 0, & r\neq s.\end{cases}$$

Gain Equation

From Eq.(2.7) we have

$$E\{[f(m,n)-\hat{f}(m,n)]e(k,l)\}=E\{[\tilde{f}^{-}(m,n)-K(m,n)e(m,n)]e(k,l)\}=0, \tag{2.15}$$

thus

$$E\{\tilde{f}^{-}(m,n)e(k,l)\}=E\{K(m,n)e(m,n)e(k,l)\}=K(m,n)R_e(m,n)\delta_{mk}\delta_{nl}. \tag{2.16}$$

For the case when m = k, n = l we have

$$K(m,n) = E\{\tilde{f}^{-}(m,n)e(m,n)\}R_e^{-1}(m,n), \tag{2.17}$$

where

$$E\{\tilde{f}^{-}(m,n)v(m,n)\} = 0, \tag{2.18}$$

and

$$e(m,n) = \tilde{f}^{-}(m,n)+v(m,n).$$

The Kalman gain can now be written as

$$K(m,n) = V_f^-(m,n;m,n)R_e^{-1}(m,n). \tag{2.19}$$

A Priori Variance Algorithm

In equation (2.14) we have defined the à priori variance, in the following we shall express the à priori variance in terms of estimate error $\tilde{f}$.

Recall that $\tilde{f}^-(m,n)=f(m,n)-\hat{f}^-(m,n)$, and from the image model (2.5) we have

$$f(m,n) = \sum_{(i,j)\in\Omega} a(i,j)f(m-i,n-j)+w_s(m,n). \tag{2.20}$$

Remember that w_s is a white random process with the variance σ_{ws}^2 thus

$$V_f^-(m,n;m,n) = \sum_{(k,l)\in\Omega}\sum_{(i,j)\in\Omega} a(i,j)E[\tilde{f}(m-i,n-j)\tilde{f}(m-k,n-l)]a(k,l)+\sigma_{ws}^2$$

$$= \sum_{(k,l)\in\Omega}\sum_{(i,j)\in\Omega} a(i,j)V_f^-(r,s,t,u)a(k,l)+\sigma_{ws}^2, \tag{2.21}$$

where $r = m-i$, $s = n-j$, $t = m-k$, $u = n-l$; and $V_f^-(.,.)$ is called the à postériori variance which is readily available. We define

$$\tilde{f}(m,n)=f(m,n)-\hat{f}(m,n). \tag{2.22}$$

Substituting Eq.(2.10) into Eq.(2.25) yields

$$\tilde{f}(m,n)=f(m,n)-\hat{f}^{-}(m,n)-K(m,n)e(m,n), \qquad (2.23a)$$

or

$$\tilde{f}(m,n)=\tilde{f}^{-}(m,n)-K(m,n)e(m,n). \qquad (2.23b)$$

Thus

$$E\{\tilde{f}^2(m,n)\}=E\{[\tilde{f}^{-}(m,n)-K(m,n)e(m,n)]^2\}. \qquad (2.24)$$

It is known that v is white noise, and from Eqs.(2.14), (2.18) and (2.21), we obtain

$$V_f^{-}(m,n;m,n)=E\{\tilde{f}^2(m,n)\}=[1-K(m,n)]V_f^{-}(m,n;m,n). \qquad (2.25)$$

The equations (2.7), (2.19), (2.21), and (2.25) are the sequential Kalman filter for the image model (2.5) and the observation (2.3).

Initialization of the Algorithm

In order to implement the Kalman filter[3] we need to define the initial conditions. In image restoration, the image field is a finite two-dimensional array; therefore the initial conditions for the estimation algorithms are the signal estimates over the image boundary.

[3]Unlike the conventional Kalman vector form filter, the sequential Kalman filter is performed pixelwise.

It is known that the current estimate of the Kalman filter is optimal so long as the previous estimates are so. Further to this, the estimates of the Kalman filter are independent of their initial values, the normal estimates will asymptotically approach to optimal values even if the initial conditions are not known exactly (Anderson and Moore, 1979). However, due to the particular modeling method involved, the sequential Kalman filter is suboptimal. It is necessary to initialize the computaion with proper initial values (i.e. the initial estimates for the support Ω). Several techniques can be used for this purpose. One of these techniques is to carry out the initial restoration over the boundary by the following algorithm,

$$\hat{f}(m,n)=f_m+K_b[g(m,n)-f_m], \tag{2.26}$$

where

$$K_b=\frac{\sigma_f^2}{\sigma_f^2+\sigma_v^2}$$

and

$$\sigma_g^2=\frac{1}{A_b}\sum_{\beta\in A_b}[g(\beta)-g_m]^2,\ g_m=E\{g\},\ f_m=g_m-E\{v\},\ \sigma_f^2=\sigma_g^2+\sigma_v^2,$$

where A_b is dimension of the image field, and $\beta \triangleq (i,j)$. It can be seen that Eq.(2.26) is essentially a linear least squares estimate with the gain K_b being the variance ratio.

Indeed, Eq.(2.26) is itself a simple pixelwise restoration algorithm in the early development of restoration techniques.

2.4 The Local Spatial Statistics Measurements

The experiments conducted by Netravali and Brasada (1975) show that at sharp transitions in image intensities, the relative sensitivity of the human visual system decreases. That is, human eyes are less sensitive to noise in the sharp transition regions (e.g. edges). It is also known (Caelli and Moraglia, 1986) that edges carry the essential information revealing the details of the image. In the following an effective but simple scheme proposed by Wallis (1976) is applied to measure the spatial activity, and the Kalman filter (i.e. Eqs.(2.10),(2.22),(2.24),and (2.28)) is modified to adapt to the local statistics.

The noise statistics utilized in the Kalman filter algorithm essentially determine the gain factor for the filter. If in Eq.(2.3) the variance σ_v^2 is given as a constant, the noise content in the image is suppressed uniformally. This often results in smooth but blurred images. We can weight σ_v^2 by the local statistics such that the restoration filter will be able to preserve edges to some extent. The local variances are defined as follows,

$$\sigma_f^2(m,n)=\sigma_g^2(m,n)-\sigma_v^2, \qquad (2.27)$$

$$\bar{g}(m,n) = \frac{1}{N_b^2}\sum_{i,j}^{D} g(m-i,n-j), \tag{2.28}$$

$$\sigma_g^2(m,n)=E\{[g(m,n)-\bar{g}(m,n)]^2\}, \tag{2.29}$$

or

$$\sigma_g^2(m,n) = \frac{1}{N_b^2}\sum_{i,j}^{D} [g(i,j)-\bar{g}(m,n)]^2, \tag{2.30}$$

where D denotes the square window over which the local statistics are calculated, and N_b^2 is the window dimension. The weight factor w_t should be chosen such that

$$w_t(m,n)=\begin{cases}1, & \forall\sigma_f^2(m,n)=0,\\ c, & \forall\sigma_f^2(m,n)>0,\ 0\le c<1\end{cases} \tag{2.31}$$

which means that in flat regions, (where $\sigma_f^2=0$), $w_t\sigma_v^2=\sigma_v^2$, while in areas of high spatial activity $w_t\sigma_v^2<\sigma_v^2$. If σ_v^2 is replaced by $w_t\sigma_v^2$, the modified Kalman filter will act normally in flat regions suppressing the observation noise. However, in a region of higher spatial activity, the Kalman filter will work on the noise, say, v_r, with reduced variance $w_t\sigma_v^2(<\sigma_v^2)$. In doing so, we allow more noise passing through the filter in high spatial activity regions.

We propose the following weighting function

$$w_t(m,n)=\exp\{-K_1[\sigma_t^2(m,n)]\}, \qquad (2.32)$$

where $K_1>0$ is a constant to be decided by experiments. It is interesting to note that Eq.(2.32) has the form similar to that of the so-called relative noise visibility function (Anderson and Netravali, 1976). However, Eq.(2.32) is obtained from the mere consideration of edge preservation and is not based on any subjective tests. For most noisy images the use of w_t alone will be sufficient, since, as stated earlier, human visual system is less sensitive to noise in areas of high spatial activity than in flat regions. It is, however, more desirable in some cases to smooth out the noise around the edge.

As we discussed before, local statistics (namely, local mean values and local variance) are calculated over a window of N_D^2 pixels with f at the window's geometric centre. For the case when an edge is encountered in the window region (see Figure 2.6), the local statistics information may be more precisely estimated, provided we could distinguish on which side of the edge the pixel f is located and determine the orientation of the edge. The Kalman filter will then be effective even along the edge. In the case of Figure 2.6, pixel f is located in the unshaded area which is a subset of the window D. Our task is thus to determine the subset and calculate the local statistics over the subset. This procedure is called refined local statistics measurement (Lee, 1981).

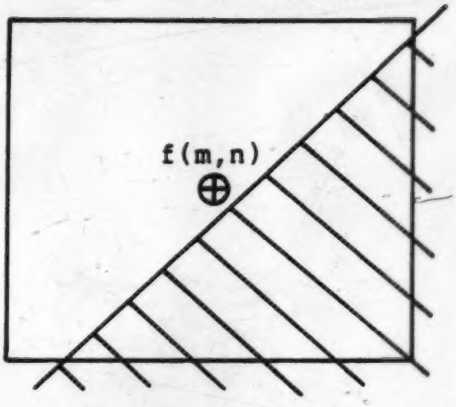


Figure 2.6 An edge across the window region

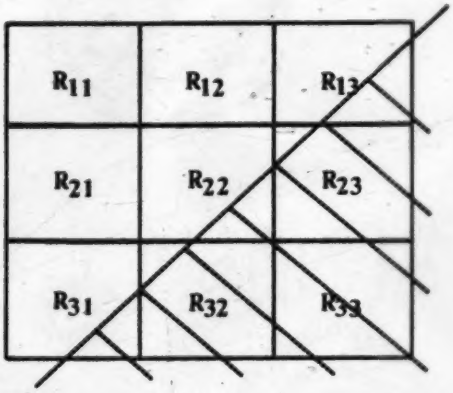
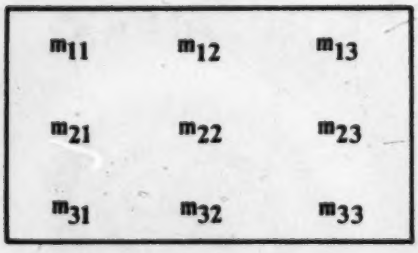
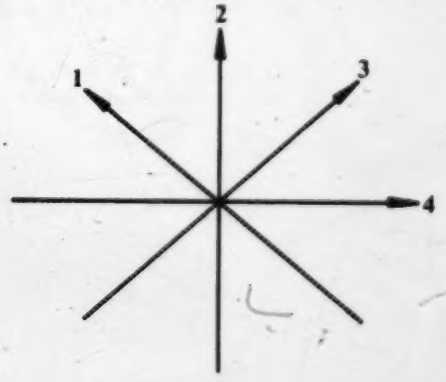


Figure 2.7 The window divided into subregions $R_{i,j}$



(a)



(b)

Figure 2.8 (a) Local means for the subregions in the window
(b) edge orientaion map

In order to determine the subset it is necessary to identify the edge orientation. This can be done by calculating mean value for each of the subregions R_{ij} as shown in Figure 2.7 where the window D is divided into 9 subregions, R_{22} is the region in which f is located, and m_{ij} denotes the mean value of R_{ij} (see Figure 2.8a). In Figure 2.8b we have also depicted four possible orientations. Since each direction has its complement, these in effect, represent four pairs of directions.

After $\{m_{ij}\}$ have been computed, a set of simple compass gradient masks is convolved with the mean values to determine the edge orientations. The gradient masks are shown below

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

It is to be noted that these are consistent with Figure 2.8b. The mask which yields the maximum absolute gradient will be chosen as the possible edge direction. For instance, if $|m_{11}-m_{33}|+|m_{12}-m_{23}|+|m_{21}-m_{32}|$ is the maximum, then the first direction is chosen, and we are left with computing $|m_{11}-m_{22}|$ and $|m_{22}-m_{33}|$ to determine the subset to which the pixel f belongs. For example, $|m_{11}-m_{22}|<|m_{22}-m_{33}|$ indicates that f is more likely in the unshaded subset, and the local mean value f and variance σ_f^2 are defined over that

subset.

2.5 Identification of the Modeling Coefficients

In the image model (2.5), the modeling coefficients are normally not known, and must be estimated. We may write Eq.(2.5) in vector form, as,

$$f(m,n)=\mathbf{A}^T(m,n)\mathbf{F}(m,n)+w_s(m,n), \tag{2.33}$$

where $\mathbf{A}$ [3] is the vector composed of the modeling coefficients, and $\mathbf{F}$ the signal vector. The structures of $\mathbf{A}$ and $\mathbf{F}$ depend on specific scanning operation and neighbourhood configuration. Eq.(2.5) is causal, and the original image is also assumed to be an ergodic random field. We may use the identification algorithm (Ljung et al., 1983) to estimate $\{a(i,j)\}$. In order to make the notations simple, we define $k \triangleq (m,n)$ and use $k+j$ to represent moving j steps further from (m,n) or k. With this definition, Eq.(2.33) can be written as

$$f(k)=\mathbf{A}(k)\mathbf{F}(k)+w_s(k). \tag{2.34}$$

From (2.34), a general parameter identification algorithm is readily available as follows assuming that the modeling coefficients $\{a(i,j)\}$ are constants,

[3] The bold face letters represent vector or matrix variables.

$$\hat{\mathbf{A}}(k+1|k+1) = \hat{\mathbf{A}}(k|k)+\Phi(k+1)[\mathbf{f}(k)-\mathbf{F}(k)\hat{\mathbf{A}}(k)], \tag{2.35a}$$

$$\Phi(k+1) = \mathbf{R}(k+1|k)\mathbf{F}^{T}(k)[\mathbf{F}(k)\mathbf{R}(k+1|k)\mathbf{F}^{T}(k)+\sigma_{vs}^{2}(k+1)\mathbf{I}]^{-1}, \tag{2.35b}$$

$$\mathbf{R}(k+1|k) = \mathbf{R}(k|k)+\sigma_{vs}^{2}(k)\mathbf{I}, \tag{2.35c}$$

$$\mathbf{R}(k+1|k+1) = [\mathbf{I}-\Phi(k+1)\mathbf{F}(k)]\mathbf{R}(k+1|k), \tag{2.35d}$$

where

$$\mathbf{R}(k|k) = E\{[\mathbf{A}(k)-\hat{\mathbf{A}}(k|k)][\mathbf{A}(k)-\hat{\mathbf{A}}(k|k)]^{T}\}, \tag{2.36}$$

and σ_{vs}^{2} is constant for stationary white process.

From Eq.(2.35) we can see that this algorithm actually uses the original image to estimate $\mathbf{A}(k)$. Unfortunately, in restoration problems, the original image is not available at all, the direct use of Eq.(2.35) is seemingly meaningless. However, in many practical applications the gross structure of the original image may be observed, and we may apply Eq.(2.35) over images with similar structure to obtain $\{a(i,j)\}$.

In most cases we may have or assume some knowledge about original images, e.g. the statistical characteristics of the images. Experiments have shown that many images can be regarded as wide-sense stationary random fields with the following autocorrelation function (Franks,1966),

$$R(i,j) = \sigma_s^2 \exp[-\alpha|i|-\beta|j|], \tag{2.37}$$

where σ_s^2 is the global image variance, i and j are the horizontal and vertical displacements, respectively, α and β are constants.

With Eq.(2.37) given we are now in a position to introduce another parameter estimation method which is more realistic than Eq.(2.35). From Eq.(2.33) we have

$$w_s(m,n) = f(m,n)-\mathbf{A}(m,n)\mathbf{F}(m,n), \tag{2.38}$$

we seek $\mathbf{A}$ such that

$$J(w_s(m,n)) = E\{w_s^2(m,n)\}, \tag{2.39}$$

is minimized. This gives

$$\mathbf{A}(m,n) = \left[\sum_{\Omega}\mathbf{F}(m,n)\mathbf{F}^T(m,n)\right]^{-1}\left[\sum_{\Omega}f(m,n)\mathbf{F}(m,n)\right]. \tag{2.40}$$

The right side of Eq.(2.40) is, in fact, composed of the autocorrelation $\{R(i,j);(i,j)\in\Omega(m,n)\}$. Since Eq.(2.37) is known Eq.(2.40) can be solved.

2.6 Simulations

In this section, several images are used to test the proposed sequential Kalman filter.

We use the following image model which takes the four closest pixels as neighbours (see Figure 2.5). The image model (2.5) can be written explicitly as

$$f(m,n) = a_1 f(m,n-1)+a_2 f(m-1,n-1)+a_3 f(m-1,n) + a_4 f(m+1,n-1)+w_s(m,n). \qquad (2.41)$$

The experiment is carried out through four steps as follows,

(1) Estimating the modeling coefficients $\{a_i\}$

From Eq.(2.41) the **A** and **F** vectors can be written explicitly,

$$\mathbf{A} = \left[a_1\ a_2\ a_3\ a_4\right],$$

$$\mathbf{F}^T = \left[f(m,n-1)\ f(m-1,n-1)\ f(m-1,n)\ f(m+1,n-1)\right].$$

The random field that represents the spatial activity of the original image is also assumed to be wide-sense stationary in both directions with the exponential autocorrelation function Eq.(2.37). From Eq.(2.40) and Eq.(2.37) we have the matrix equation

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} R_{00} & R_{10} & R_{11} & R_{10} \\ R_{10} & R_{00} & R_{01} & R_{20} \\ R_{11} & R_{01} & R_{00} & R_{21} \\ R_{10} & R_{20} & R_{21} & R_{00} \end{bmatrix}^{-1} \begin{bmatrix} R_{01} \\ R_{11} \\ R_{10} \\ R_{11} \end{bmatrix}, \tag{2.42}$$

where $R_{ij} \triangleq R(i,j)$.

(2) Choosing the constant K_1 for the weighting function. In the experiments to follow, several different values of K_1 were used (ranging from 150 to 600 for the Girl image, for instance) in each case. The final values of K_1 are determined based on the subjective judgment (K_1=370 is used as the value for the Girl image, for example).

(3) Estimating the observed image using Eqs.(2.7), (2.19), (2.21), and (2.25)

(4) Preserving edges using the refined local statistics.

Figure 2.9a shows a Girl image having a 256x256 format. Figure 2.9b is the Girl image corrupted by Gaussian white noise with standard deviation σ_v^2=400 (relative to 8-bit resolution), and zero mean. Figure 2.9c shows the restored image using Eqs.(2.7), (2.19), (2.21), and (2.25) only. The over all image is smooth but blurred. Figure 2.9d is the image restored by using Eqs.(2.7), (2.19), (2.21), (2.25) and the weighting function $w_t(m,n)$ with K_1=370. The refined edge preserving procedure is performed only when

$\sigma_t^2(m,n) > \sigma_v^2 + 50$. Compared with Figure 2.9c, improvement on contrast and edge details can be easily observed.

In Figures 2.10 shows W.C. Fields, letter F, and square pattern images and their histograms.[4] All the images are of 128x128 8-bit format.

In order to test the effectiveness of the proposed Kalman filtering procedure, these images are corrupted by Gaussian noise with zero mean values but different variances. Figures 2.11(a), 2.12(a), and 2.13(a) show the images contaminated by additive Gaussian noise with $\sigma_v^2=325$, Figures 2.11(b), 2.12(b), and 2.13(b) are the images plus Gaussian noise with $\sigma_v^2=900$, and the images in Figures 2.11(c), 2.12(c), and 2.13(c) are the noisy images with $\sigma_v^2=1600$. The restoration results using Eqs.(2.7), (2.19), (2.21), (2.25) and the refined edge preserving method are shown in Figures 2.11-2.13(d), (e), (f) (the second row). Figures 2.11-2.13(g), (h), (i) (the third row) are the restored images using Eqs.(2.7), (2.19), (2.21), (2.25) and the weighting function with $K_1=350$, 820, and 1550, respectively. All the restored images show sharp edges and the noise is suppressed as well.

---

[4] In the image processing literature, histogram is defined over an image. Given an image I, let $F_I(c)$ be the relative frequency with which grey level c occurs in I, for all c in the grey level range $[c_1,c_t]$ of I. The graph of $F_I(c)$ as a function of c is called the histogram of I.

Figure 2.9 The Girl image.
(a) The original 256x256 digital image. (b) The original image degraded with additive noise v with variance σ_v^2=400. (c) Restored image of the degraded image of (b) without using the weighting function w_t in the filter. (d) Restored image of the degraded image of (b) using the weighting function w_t in the sequential Kalman filter.

Pictures (Figure 2.9) are low in resolution due to the quality of the plotter, and are excluded micro-filming.

Figure 2.10 The test images and their histograms. The first row (from top): the W.C. Fields image (left) and its histogram. The second row: the letter F image (left) and its histogram (right). The third row: the square pattern (left) and its histogram (right). All the images are of 128x128 and 8-bit in resolusion.

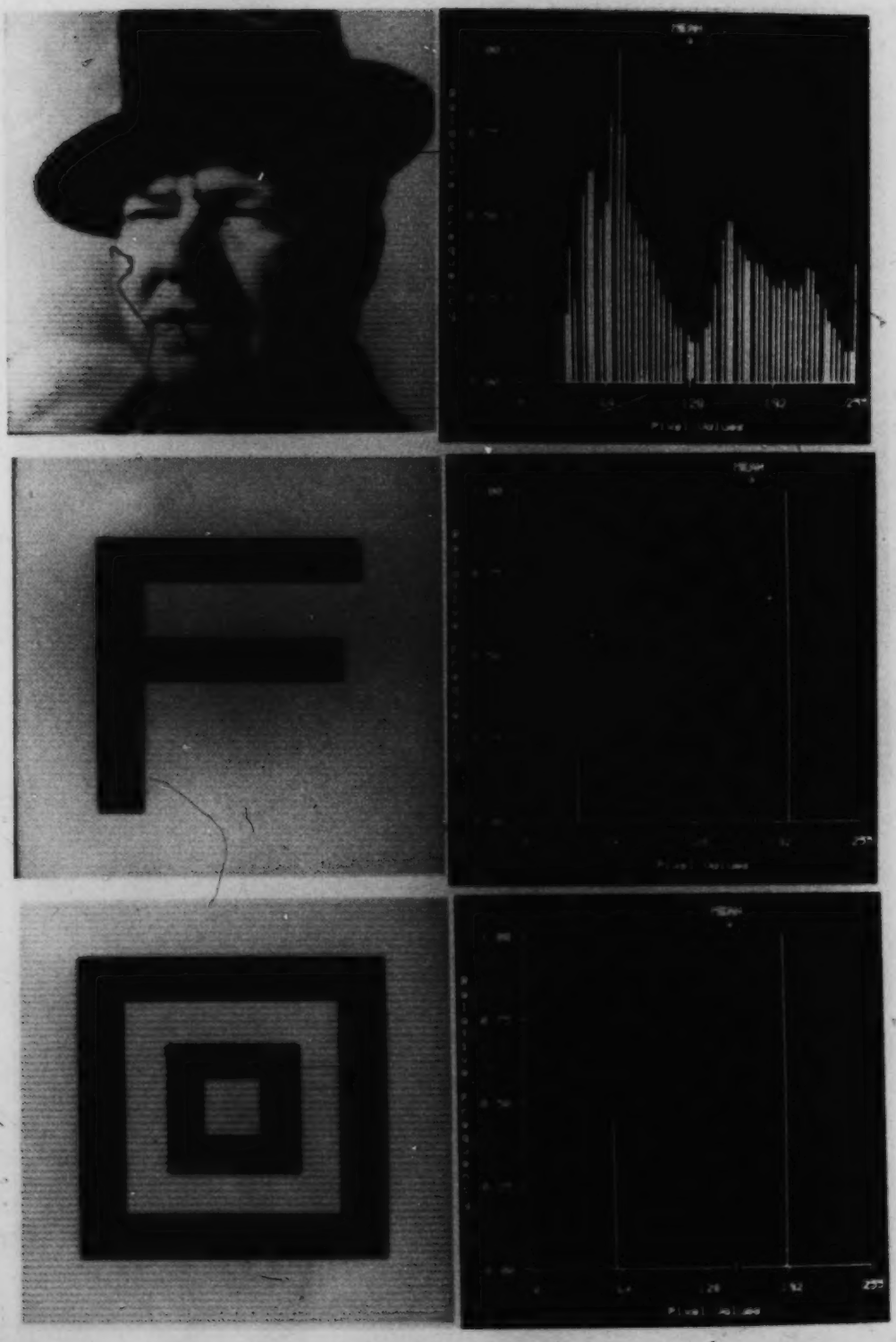


Figure 2.11 The noisy and restored W.C. Fields images. The first row is the W.C. Fields image degraded with additive noise with variances σ_v^2=325, σ_v^2=900, and σ_v^2=1600, respectively (from left to right). The second row is the restored images of the first row using the weighting function and the refined spatial activity measurement. The third row is the restored images of the first row using the weighting function w_t. All images are of 128x128 and 8-bit in resolution.

Figure 2.12 The noisy and restored letter F images. The first row is the letter F image degraded with additive noise with variances σ_v^2=325, σ_v^2=900, and σ_v^2=1600, respectively (from left to right). The second row is the restored images of the first row using the weighting function and the refined spatial activity measurement. The third row is the restored images of the first row using the weighting function w_t. All images are of 128x128 and 8-bit in resolution.

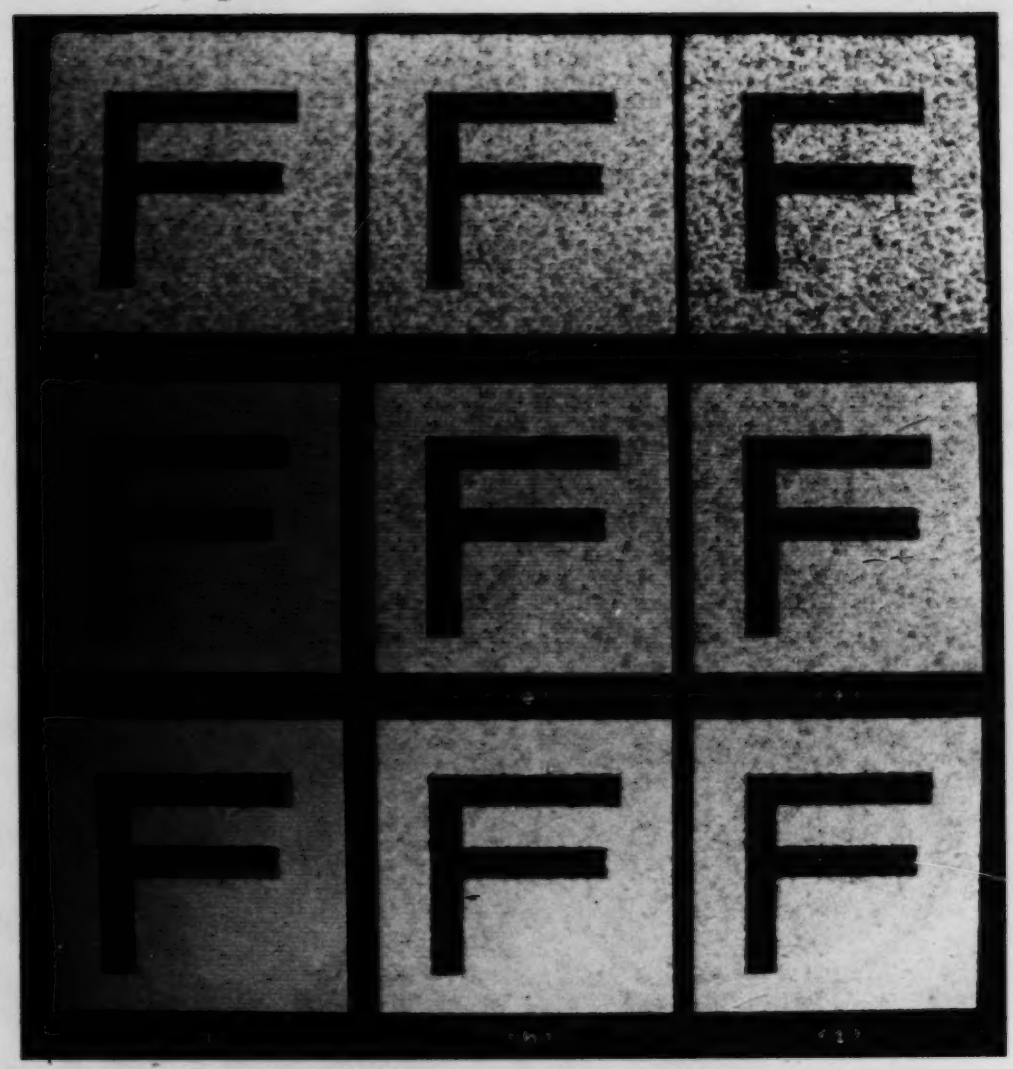
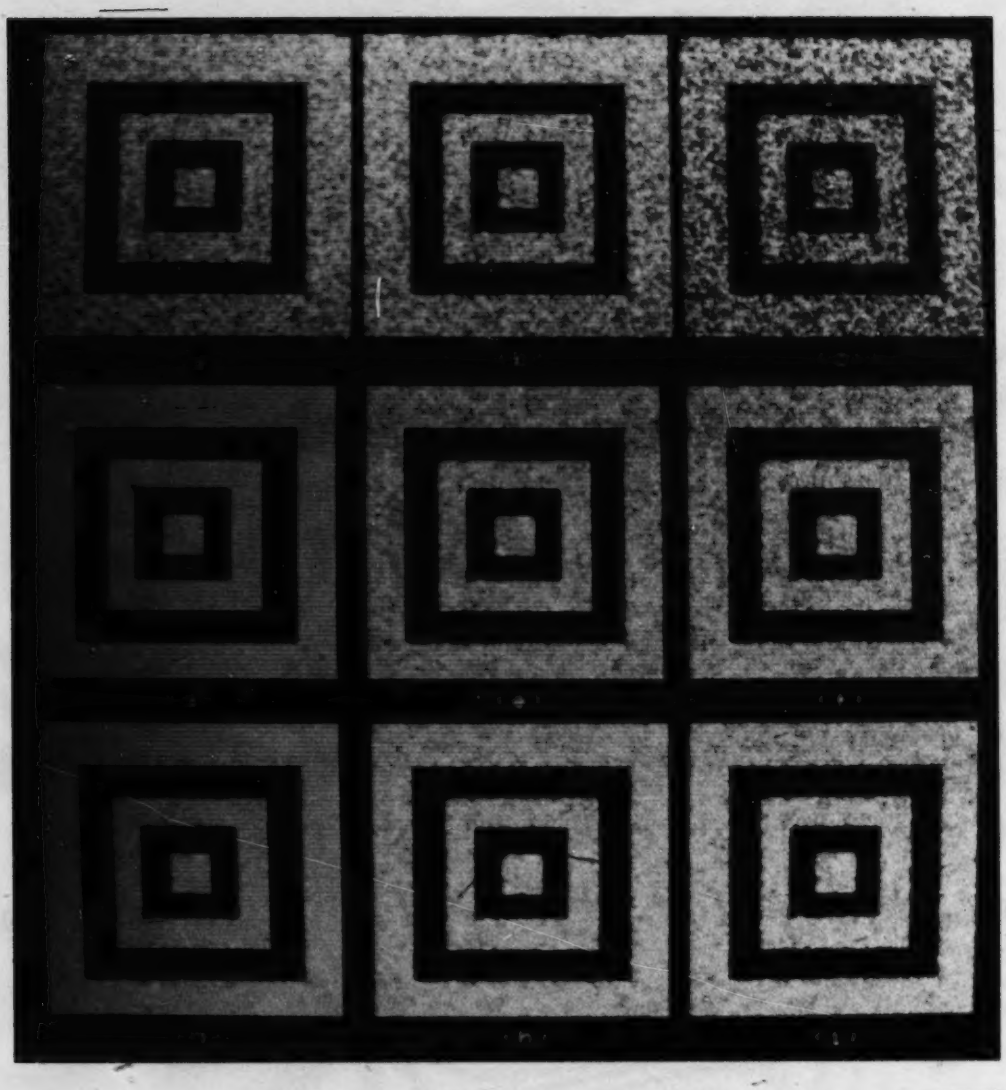

F F F
F F
F F F F

Figure 2.13 The noisy and restored square patterns. The first row shows the square pattern image degraded with additive noise with variances σ_v^2=325, σ_v^2=900, and σ_v^2=1600, respectively (from left to right). The second row is the restored images of the first row using the weighting function and the refined spatial activity measurement. The third row is the restored images of the first row using the weighting function w_t. All images are of 128x128 and 8-bit in resolution.



CHAPTER III

ADAPTIVE ITERATIVE IMAGE RESTORATION ALGORITHM

3.1 Introduction

In the previous chapter a sequential adaptive Kalman filter was developed. The derived Kalman filter were based on a causal image model. However, for image restoration we are given a complete image field, causality of the image field is artificially contrived to satisfy the mathematical requirements of the Kalman filter. Since we have all the information of an image field, we may develop some algorithms which resemble the structure of the Kalman filter but based on a noncausal image model. In this chapter we shall develop an adaptive algorithm which is implemented iteratively for image restoration.

The main objective of image restoration is to restore an original image given noisy observations of it and à priori knowledge about the image and noise statistics. Problems of this kind in the signal estimation literature are termed inverse problems, since the task is to infer the input from the output.

Early research on image restoration techniques was mainly conducted in the frequency domain where computations of nonrecursive algorithms (e.g. the Wiener filter) were carried out using fast Fourier transform (FFT) techniques.

But the inherent limitation of these algorithms is that the image is required to be a wide-sense stationary field. The restoration algorithms based on this assumption, though using the computational advantages of the FFT, are insensitive to abrupt changes in the image intensity, and, as a result, tend to smooth out edges where stationarity is not justified. Another cause of this smoothing effect is the minimum mean squared error (MMSE) type criteria which are often used in the signal estimation practice and are the foundation of many estimation theories. The resulting algorithms tend to suppress noise in the low frequency range of the image rather than higher frequency components although it is the latter which, on a global basis, contain more of the image variance or entropy. This has the effect of producing essentially low-pass restoration results, and prohibiting the restoration of the input image's sharpness, particularly, in the edge areas.

Consequently, with a stationary image field (shift-invariant), these techniques have error components which are amplified as a function of local contrasts, and particularly in edge regions. However, it is known that the human visual system (configured with detectors whose point spread functions or receptive fields closely approximate band-pass filters (Caelli and Moraglia, 1986)) is highly sensitive to edge information and to regions of high contrast change. This means that the human eye is less sensitive to noise in the high-pass regions of the image

frequency components and more sensitive to noise in these low-pass regions. It is for these reasons that the stationary and MMSE methods do not produce restoration solutions which are visually satisfying.

This observation has led to a number of restoration algorithms that include the human viewer as a factor. One of more obvious developments is due to Anderson and Netravali (1976). Instead of using a nonstationary image model, they used a subjective error criterion based on human visual models and derived a nonrecursive filter that adapts itself to make a compromise between loss of sharpness of resolution and smoothness of noise such that the same amount of subjective noise is suppressed throughout the image. Later, Abramatic and Silverman (1982) generalized this procedure and incorporated it in the classic Wiener filter.

Perhaps Hunt and Cannon (1976) were the first to seriously tackle nonstationary models for image fields. They demonstrated that images are, in general, nonstationary. In order to make use of the Gaussian probability density function (PDF), they assume that the means of image pixels are nonstationary, while the variances are stationary. With this Gaussian model they developed a maximum à postériori (MAP) restoration algorithm (Hunt, 1977). Further to this, Kuan et al. (1984) proposed a nonstationary mean, nonstationary variance (NMNV) statistics model for image fields. As in the Hunt and Cannon's model, they use a

nonstationary mean to describe the gross structure of an image. The difference is that the variance which characterizes edge and elementary texture information of the image is nonstationary. More recently, Kasturi et al. (1985) have reported an algorithm to restore images corrupted by signal dependent noise using the MMSE criterion. They assume that the image is statistically governed by the Gaussian PDF. Many other iterative and recursive algorithms with various methods to handle the nonstationarity of the image have also been developed recently along this line (e.g. Trussell and Hunt, (1978, 1979); Dudgeon, (1980); Schafer et al., (1981); Yamakoshi et al., (1982)).

As already mentioned, most algorithms using the MMSE or LMMSE (Linear MMSE) criteria result in globally smoothed images. However, as is evident from the above discussion, these criteria should be used locally to improve the visual quality of the restored image.

It is the author's impression that, in many if not most of the recent nonstationary restoration algorithms, the PDF of the image field is assumed to be Gaussian for the convenience of mathematical derivations, and much of the effort is devoted to modeling the image field. Little attention, if any, is given to the observation model. In this chapter we propose a new observed image model which takes the local statistics into account, and a noncausal model of the original image containing unknown parameters.

In particular, we do not assume that the image and the noise statistics are Gaussian. The algorithm is constrained by a linear mean variance criterion operating within local regions of the image. That is, the image is decomposed into different sections according to spatial statistics of the image. The algorithms are derived for each of these regions. Further, the nonstationarity of the image is handled after the restoration solution is generated. The proposed observed image model enables the local factor to be automatically included in the algorithm gain equation. This factor, which varies with the local statistics, in turn adapts the gain to the image local contrast and luminance.

3.2 Image Modeling

In the previous chapter we established a causal configuration which assumes that any image of interest can be decomposed, depending on the location of the pixel being processed, into three parts, namely, "past", "present", and "future". This enables one to establish an autoregressive image model which linearly relates a given pixel to a specified set of its neighboring pixels in the "past". Obviously, the use of such linear representations for the image involves uncertainty. In order to take this into account, a white process w_s called modeling error is introduced. We pointed out in the Introduction that in image restoration a complete image field is available. Consequently, we can establish a noncausal configuration: A

pixel is related to its neighboring pixels in both "past" and "future" regions. More formally, we denote a rectangular region surrounding the pixel f(m,n) by A_s as shown in Figure 3.1 being defined by

$$A_s \triangleq [(m,n):\{-r_1 \le m \le r_2, -s_1 \le n \le s_2, \ \forall (m,n) \ne (0,0)\}]. \qquad (3.1)$$

The image model is given as follows

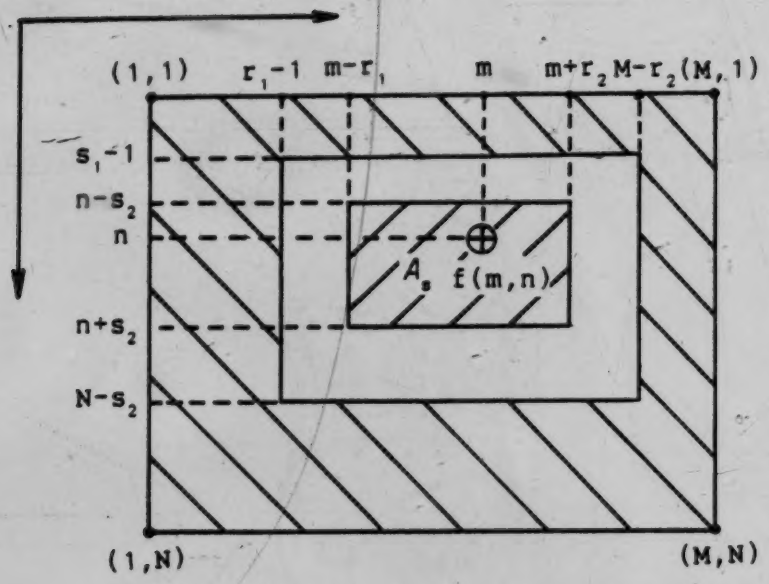
$$f(m,n) = \sum_{(i,j)\in A_s} \theta(i,j) f(m-i,n-j) + w(m,n), \qquad (3.2)$$

where $\theta(.,.)$ are the modeling coefficients, and w(.,.) is the modeling error which may be assumed to be a random process. Recall that in Eq.(2.2) we assumed that w_s to be a white process. For a noncausal image model, however, it is shown that w is nonwhite (Biemond et al., 1983). For a discrete image field, Eq.(3.2) can be written in a compact vector form as follows

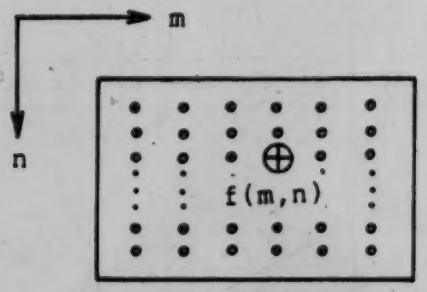
$$f(m,n) = \Theta^T F(m,n) + w(m,n). \qquad (3.3)$$

The linear observed image model is, in general, has the following form,

$$g(m,n) = H(m,n) * f(m,n) + v(m,n), \qquad (3.4)$$



(a) The Image Field



(b) The Window Area A_s

Figure 3.1 Noncausal image field

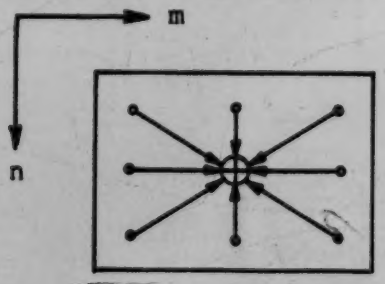


Figure 3.2 The window A_s used in the simulations

where * denotes convolution, and H is the point spread function (PSF) of imaging systems.

In the development of various restoration algorithms concerning the suppression of noise, the statistical characteristics of v have great impact on the resulting algorithms. For instance, in the Kalman filter technique, the gain K is governed by the noise variance. In the MAP estimator (Hunt, 1977) which assumes the noise PDF to be Gaussian, the estimate is

$$\hat{\mathbf{f}} = (\mathbf{R}_f^{-1}+\mathbf{R}_v^{-1})^{-1}\mathbf{R}_v^{-1}\mathbf{g}+(\mathbf{R}_f^{-1}+\mathbf{R}_v^{-1})^{-1}\mathbf{R}_f^{-1}\mathbf{f}_m, \tag{3.5}$$

where bold letters denote vector forms of the corresponding functions, $\mathbf{f}_m=E\{\mathbf{f}\}$, and $\mathbf{R}_v$ and $\mathbf{R}_f$ are the covariance matrices of v and f respectively. We see that $\hat{\mathbf{f}}=\mathbf{f}(\mathbf{R}_v,\mathbf{R}_f)$. These are natural and expected results. Since the purpose is to estimate images from noisy environments, the noise characteristics are essential to the development of these algorithms. Here we attempt to use the local spatial activity of the image as a part of the restoration algorithm, so that the image model is still assumed to be stationary and the PDF of the noise is not necessarily known. To do this, we first introduce a rather simple modification to the observed image model (3.4). The new observed image model is

$$g_1(m,n) = f(m,n)+v(m,n)-\eta(m,n)v(m,n)$$

or

$$g_1(m,n) = f(m,n)+[1-\eta(m,n)]v(m,n), \tag{3.6}$$

note that we have assumed H to be an identity (delta PSF), where $\eta(m,n)$ is a variable which takes local statistical information into account. The variable is defined by

$$\eta(m,n)= \frac{\sigma_f^2(m,n)}{\sigma_f^2(m,n)+\sigma_v^2(m,n)}, \tag{3.7}$$

where $\sigma_f^2(m,n)$ is the à priori variance of f(m,n), defined as

$$\sigma_f^2(m,n) = \begin{cases} \sigma_g^2(m,n)-\sigma_v^2(m,n), & \forall \sigma_g^2(m,n) \geq \sigma_v^2(m,n), \\ 0, & \text{otherwise}, \end{cases} \tag{3.8}$$

where $\sigma_g^2(m,n)$ is the à priori variance of g(m,n), and $\sigma_v^2(m,n)$ is the noise variance. In a window D of size (2K+1)x(2L+1), $\sigma_g^2(m,n)$ is given by the approximation:

$$\sigma_g^2(m,n) = \frac{1}{(2K+1)(2L+1)} \sum_{i,j}^{D} [g(i,j)-\bar{g}(m,n)]^2, \tag{3.9}$$

where $\bar{g}(m,n)$ denotes the local sample average of g(m,n) in the window D and is approximated by

$$\bar{g}(m,n) = \frac{1}{(2K+1)(2L+1)} \sum_{i,j}^{D} g(m-i,n-j). \tag{3.10}$$

Substituting Eq.(3.8) into Eq.(3.7) yields

$$\eta(m,n) = \begin{cases} 1 - \dfrac{\sigma_v^2(m,n)}{\sigma_g^2(m,n)}, & \forall \sigma_g^2(m,n) \geq \sigma_v^2(m,n), \\ 0, & \text{otherwise.} \end{cases} \tag{3.11}$$

Clearly, $\eta(m,n)$ varies with the pixels involved and $0 \leq \eta \leq 1$. It will be shown that with this modified version of the observed image model, the algorithm adapts itself to local activity. Therefore, in terms of adaptation, we have a nonstationary restoration algorithm.

3.3 The Restoration Algorithm

We propose a suboptimal filter of the form

$$\hat{f}_i(m,n) = \hat{f}_i^-(m,n) + K_i[g_1(m,n) - \hat{f}_i^-(m,n)], \tag{3.12}$$

where $\hat{f}_i^-(m,n)$ is the predicted estimate of $f(m,n)$ at the i-th iteration. This filter includes two parts; a prediction part, and an updating part which improves the estimate based on the new observation $g_1(m,n)$. The predicted estimate after the i-th iteration is given by

$$\hat{f}_i^-(m,n) = \hat{\theta}_i^T \hat{F}_{i-1}(m,n), \tag{3.13}$$

where $\hat{\theta}_i$ is the i-th estimate of the coefficient vector θ, and $\hat{F}_{i-1}$ is the estimate of the vector F during the (i-1)th iteration. Combining equations (3.12) and (3.13), we get

$$\hat{f}_i(m,n) = (1-K_i)\hat{\theta}_i^T \hat{F}_{i-1}(m,n)+K_i g_1(m,n). \tag{3.14}$$

It is often the case that, in practical situations, one has little or no knowledge of the original image model. Therefore, it is necessary to invoke an identification procedure for estimating the parameter vector θ. At the i-th iteration, we define

$$J(\hat{\theta}_i) = \sum_{\Omega}\left[\hat{f}_{i-1}(m,n)-\hat{\theta}_i^T \hat{F}_{i-1}(m,n)\right]^2, \tag{3.15}$$

and

$$J(K_i) = E\{[\hat{f}_i(m,n)-f(m,n)]^2\}, \tag{3.16}$$

where Ω may be the entire image field or a window in it. Clearly, Eq.(3.15) is a least squares error criterion which is minimized for certain $\hat{\theta}_i$, while Eq.(3.16) is a mean squares error criterion. These two criteria are the most common forms used in system identification and estimation procedures.

Minimizing Eq.(3.15) with respect to the parameter vector $\hat{\theta}_i$, we have

$$\sum_{\Omega}\left[\hat{F}_{i-1}\hat{F}_{i-1}^T\hat{\theta}_i-\hat{f}_{i-1}\hat{F}_{i-1}\right] = 0, \tag{3.17}$$

which gives

$$\hat{\theta}_i = \left[\sum_{\Omega} \hat{F}_{i-1}\hat{F}_{i-1}^T\right]^{-1}\left[\sum_{\Omega} \hat{f}_{i-1}\hat{F}_{i-1}\right], \qquad (3.18)$$

where we have dropped (m,n) for notational simplicity.

To compute the filter gain K_i from Eq.(3.16), we assume ergodicity and that $\theta_i^T F_{i-1}$ and v(m,n) are uncorrelated. Substituting Eq.(3.14) into Eq.(3.16), we get

$$J(K_i) = \frac{(1-K_i)^2}{(2P+1)(2Q+1)} \sum_{\Omega}\left[g_1(m,n)-\hat{\theta}_i^T \hat{F}_{i-1}\right]^2+\left[2K_i-1\right]\sigma_v^2(m,n), \qquad (3.19)$$

where (2P+1)x(2Q+1) is the window Ω dimension which may be the entire image.

Minimization of Eq.(3.19) with respect to K_i yields

$$K_i = 1-\frac{\left[1-\eta(m,n)\right]N_\Omega\sigma_v^2}{\sum\left[\hat{\theta}_i^T \hat{F}_{i-1}-g_1(m,n)\right]^2}, \qquad (3.20)$$

where $N_\Omega=(2P+1)(2Q+1)$. In an actual application g_1 is replaced by g, since only the latter is available. Equations (3.18),(3.20) and (3.14) constitute the restoration algorithm. To initialize the computation we simply set $\hat{f}_0=g(m,n)$.

We know that η varies from 0 to 1 depending upon local spatial activity of the image. Also, from Eq.(3.20), it is

clear that the gain K_i varies from 0 to some value less than 1. In a flat region where the signal luminance distribution is dominated by additive noise, η=0 and Eq.(3.6) is identical to Eq.(3.4). On the other hand, in a region where edge is encountered, the image signal variance $\sigma_f^2(m,n)$ is much higher than that of the noise, thus η=1 making K_i=1. From Eq.(3.14) we obtain the estimated $\hat{f}_i$=g which is the noisy observed image itself.

3.4 Simulations

In order to evaluate the effectiveness of the proposed observed image model and algorithm, we have examined several different images. In this simulation the image model has the following explicit form (see Figure 3.2),

$$\begin{aligned} f(m,n) = {} & \theta_1[f(m-1,n-1)+f(m+1,n+1)] \\ & +\theta_2[f(m-1,n)+f(m+1,n)] \\ & +\theta_3[f(m,n-1)+f(m,n+1)] \\ & +\theta_4[f(m-1,n+1)+f(m+1,n-1)]+w(m,n). \end{aligned} \tag{3.21}$$

This can be conveniently written in a compact vector form as follows

$$f(m,n) = \Theta^T F(m,n)+w(m,n), \tag{3.22}$$

where

$$\Theta=\begin{bmatrix}\theta_1\\ \theta_2\\ \theta_3\\ \theta_4\end{bmatrix}, \qquad F(m,n)=\begin{bmatrix} f(m-1,n-1)+f(m+1,n+1)\\ f(m-1,n)+f(m+1,n)\\ f(m,n-1)+f(m,n+1)\\ f(m-1,n+1)+f(m+1,n-1)\end{bmatrix}$$

Figure 3.3a shows a facial image of size 256x256, Figure 3.3b is the image degraded by a white, Gaussian noise sequence with zero mean and variance σ_v^2 = 830. Three iterations and the moving window method (with a window size of 15x15) were used. Two different window sizes were chosen in this experiment for estimating the local statistics. Figure 3.3c shows the processed image using Eq.(3.22), Eq.(3.6) and the algorithm by setting η to 0. When compared with Figure 3.3b, we can see that the noise is removed and the image is fairly smooth. Figure 3.3d (for the case where W=7x7) and Figure 3.3e (for the case where W=9x9) show the results obtained by using the algorithm. It is apparent that the edges are preserved and the processed images retain some details and are less blurred, though not as smooth as Figure 3.3c.

To further reveal the edge preserving property of the proposed restoration algorithm, an artificial pattern image of size 260x260 was generated (Figure 3.4a). This image, contaminated by additive Gaussian noise with zero mean and

variance of σ_v^2=1225 is shown in Figure 3.4b. The image in Figure 3.4c is the image of Figure 3.4b restored by using the algorithm with η set to 0. Figure 3.4d is the image of Figure 3.4b restored by using the algorithm proposed in this chapter. Here a window size of W=7x7 is used. Comparing Figure 3.4c and Figure 3.4d, we can clearly see the blurred edges and overall smoothness of the image in Figure 3.4c. However, in Figure 3.4d, the edges are well preserved and, in flat regions, the noise is suppressed. In Figure 3.5a to Figure 3.5d the profiles of the corresponding images were plotted for comparison, the profile data were taken along the line (J,65) in the image fields.

The quantitative evaluations of the restored images are given in Table 3.1, where MSE is the Mean Squares Error, MAE the Mean Absolute Error, and SNR the Signal to Noise Ratio. It is seen that the overall error measurement for the proposed algorithm is slightly higher than that for the algorithm with η set to 0.

The three images used in Chapter II (Figure 2.10) and their noisy images (see the first row of Figures 2.11-2.13) were also tested. Figures 3.6a-3.6c (the first row) show the restored W.C. Fields images using the proposed adaptive algorithm. The restoration results by setting η to 0 are given in Figures 3.6d-3.6f (the second row). The test results for letter F and the square pattern are given in Figures 3.7 and 3.8 in the same arrangement as that of

Figure 3.6. All the results are obtained after three iterations and a window of size W=7x7. Inspection of these results shows that by setting η = 0 the restored images are smooth and blurred, while the images restored by using the adaptive restoration algorithm are considerably sharper with the preserved edges. This is to be expected in view of the features of the proposed algorithm. This algorithm adapts its gain K to local statistical measurement and less noise is suppressed around edges. It can be seen from the simulation results that by incorporating local statistical information into restoration algorithms one can obtain restored images in which the edges are preserved and which, consequently, are more appealing to the human observer.

Table 3.1 Quantitative measurements of image quality

FIG.NO.	MSE	MAE	SNR
3b	366	15.4	41
3c	106	8.04	53
3d	156	9.35	49
4b	1263	29.3	28
4c	945	12.1	31
4d	1004	13.5	30

Figure 3.3 The technician image.
(a) The original 256x256 image. (b) The original image degraded with additive noise v with variance σ_v^2=830. (c) The restored image of (b) by setting η to 0. (d) The restored image of (b) with the local spatial activity measurement η and window size w=7x7. (e) The restored image of (b) with the local spatial activity measurement η and window size w=9x9.

Pictures (Figure 3.3) are low in resolution due to the quality of the plotter, and are excluded from micro-filming.

Figure 3.4 The artificial pattern image.
(a) The original 256x256 pattern image. (b) The original image degraded with additive noise v with variance σ_v^2=1225. (c) The restored image of (b) by setting η to 0. (d) The restored image of (b) with the local spatial activity measurement η and window size w=7x7.

Pictures (Figure 3.4) are low in resolution due to the quality of the plotter, and are excluded from micro-filming.

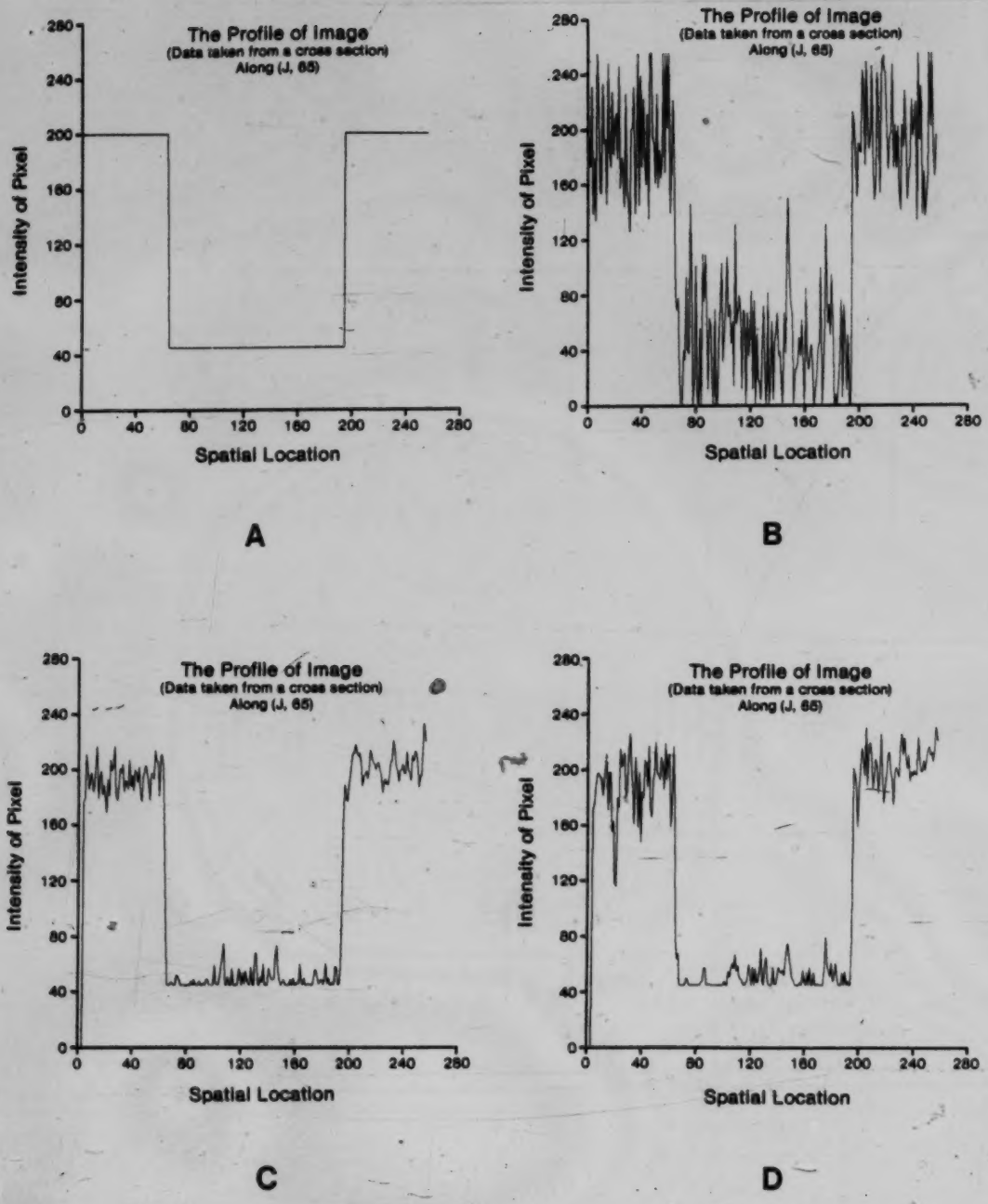


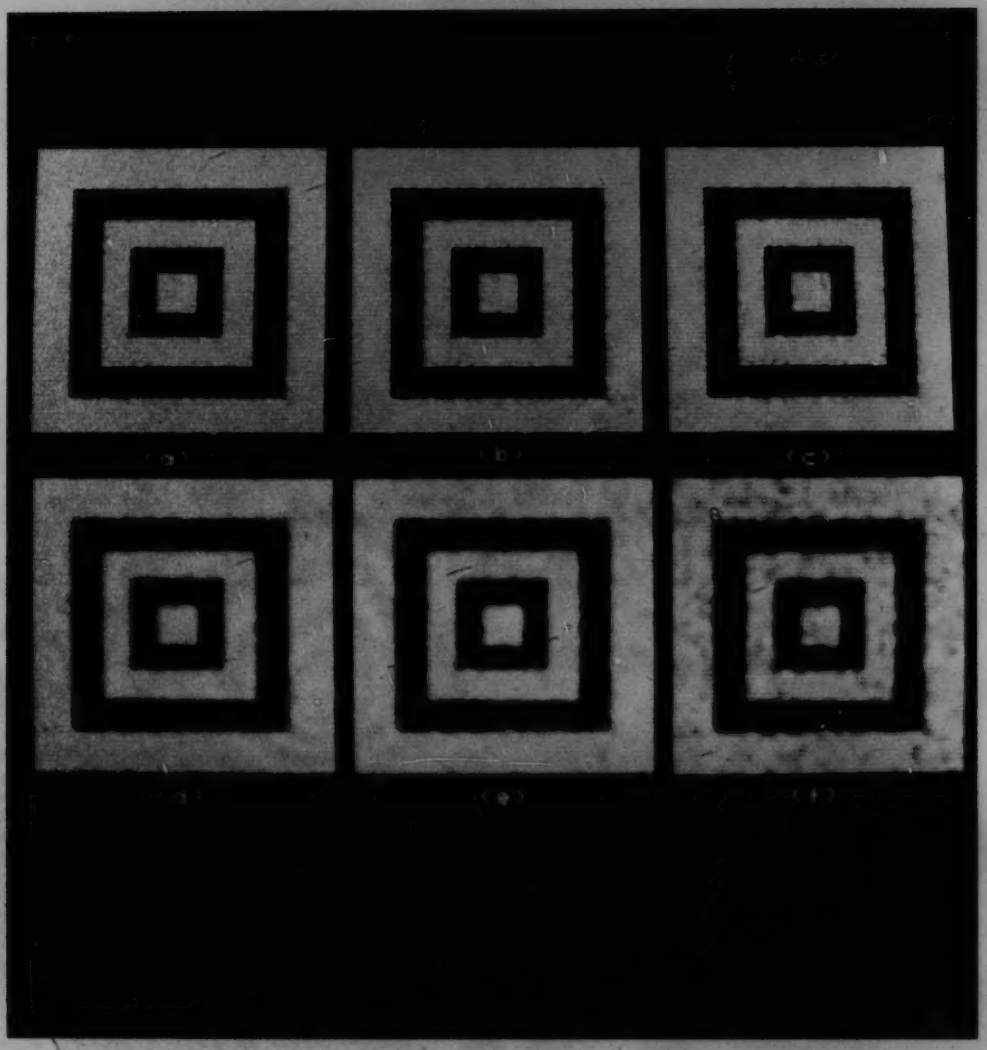
Figure 3.5 Cross section profiles of the images in Figure 3.4

Figure 3.6 Restored W.C. Fields images.
The noisy images are those in the first row of Figure 2.11.
In the first row of this figure are the restored images with
local spatial activity measurement η and w=7x7. The second
row shows the restored images by setting η to 0.

Figure 3.7 Restored letter F images.
The noisy images are those in the first row of Figure 2.12. In the first row of this figure are the restored images with local spatial activity measurement η and w=7x7. The second row shows the restored images by setting η to 0.

F F F
F F F

Figure 3.8 Restored square pattern images. The noisy images are those in the first row of Figure 2.13. In the first row of this figure are the restored images with local spatial measurement η and w=7x7. The second row shows the restored images by setting η to 0.



CHAPTER IV

AN ADAPTIVE MATCHED FILTER FOR DETECTING IMAGES IN NONWHITE NOISE BACKGROUNDS

4.1 Introduction

In the previous two chapters we have considered images corrupted by noise and obtained some solutions to their restoration. In many other areas where machines are used for recognition or detection of objects in an observed image (for instance, radar target detection, robot vision, diagnosing tumour in an X-ray photography), noise and other forms of degradations will be present. Searching for efficient and accurate filters for detecting images under these conditions has resulted in many different approaches.

Detecting the presence of objects in an observed image is a problem similar to that of restoration in the sense that both have to combat noise and other forms of degradations. The main difference is that in detection the objects to be identified are usually known à priori, whereas in restoration, the objects (the original images) are not known exactly.

As with the development of various restoration algorithms, the detectors are generated mostly from their one-dimensional counterparts. Although detecting signals has become a fully fledged research area in both communication

and information science, the concept of digital image detection is relatively recent. The main stumbling block has been the need for large-scale computing facilities.

The two-dimensional detection theory basically follows the same main stream of its one-dimensional counterparts with some variations (Van Trees, 1968; Duda and Hart, 1973). Among many reported two-dimensional detectors, the classical optimal filter based on the signal-to-noise ratio (SNR) draws the most attention due to its simplicity and efficiency. It is known that this classical optimal detector is the object image itself when the noise backgound is a stationary, white process. This leads to a simple cross-correlation operation which is called matched filtering.

Unfortunately, most realistic problems related to the detection of a digital signal embedded in an image involve matching procedures which cannot assume that the embedding background image is stationary white noise. This has recently led to the investigation of signal matching techniques not based on the "matched filter theorem" (Rosenfeld and Kak, 1982), in other words, the ideal or best way to prefilter the image does not necessarily involve the signal itself. Under the non-white circumstances, edge-only matching and matching by the combination of specific features are some examples of some of the techniques used. However, the question still remains as to the "best" type of

prefiltering to use in such conditions and to the relative efficiency of various techniques.

Notwithstanding, the foregoing matched filters are still frequently used for detection, even in nonwhite noise cases, due to their simplicity and the ease of implementation. However, this results in nonoptimal filtering or nonoptimal signal detection performance in strict theoretical sense. Recently, Vijaya Kumar and Carroll (1984) quantitatively analyzed the loss of optimality in cross-correlators (or equivalently, in matched filters) compared to prefiltering by the optimal filter in a general setting. Their results serve as guide lines for the use of cross-correlators in practice, since they were able to define its "inefficiency" or loss of optimality as a function of the embedding image. More recently, Liu and Caelli (1986) have proposed a postfiltering technique which uses an adaptive filter to improve the detectability of the matched filter in presence of nonwhite noise.

In designing the optimal filter, we need to know the second order statistics of the noise (which are normally not available), and to invert its (usually very large) covariance matrix. The latter operation is generally very time consuming though it is well known that for the special cases of stationary noise the covariance matrix is Toeplitz and the inversion is relatively simple (Rino, 1970; Arcese et al., 1970; Gray, 1971).

In this chapter we are concerned with developing a signal reference adaptive technique for determining a suboptimal postfilter which significantly improves detectability of the signal as compared to matched filters (cross-correlators).

4.2 Derivation of the Optimal Detection Filter

In general terms the signal detection problem studied here is that of detecting a signal f(x,y) embedded in an image g(x,y) defined by

$$g(x,y) = f(x,y)+v(x,y), \tag{4.1}$$

where v(x,y) is the non-white "background" image. The aim is to find a filter r(x,y) convolved with g(x,y) to result in $Z_{rg}(x,y)$ which optimizes detection. That is

$$Z_{rg}(x,y) = r(x,y)*g(x,y), \tag{4.2}$$

should maximize the SNR defined as (Thomas, 1969; Arcese, 1970),

$$SNR(x,y)= \frac{|E[Z_{rg}(x,y)]|^2}{Var[Z_{rg}(x,y)]}. \tag{4.3}$$

Without loss of generality, we may assume that the signal is

centred at (0,0), i.e.

$$SNR(x,y) = SNR(0,0). \tag{4.4}$$

The optimal form of r(x,y) (in Eq.(4.2)) to result in the best SNR(0,0) has been developed for many years and can be found in many places (Van Trees, 1968; Papoulis, 1984).

In the following we shall briefly describe the optimal filter for completeness and convenience of readers. It would be preferable to derive the filter in discrete form since it is to be implemented on a digital computer. Thus we may use the discrete observed image model

$$g(m,n) = f(m,n)+v(m,n). \tag{4.5}$$

For signal centred at (0,0), Eq.(4.2) becomes

$$Z_{rg}(0,0) = r(0,0)*g(0,0) = r(0,0)*\Big[f(0,0)+v(0,0)\Big]$$

$$= \sum_{j=1}^{N}\sum_{i=1}^{M} r(-i,-j)\Big[f(i,j)+v(i,j)\Big], \tag{4.6}$$

without loss of generality we assume $E\{v\}=0$, thus

$$E\{Z_{rg}(0,0)\} = E\{\sum_{j=1}^{N}\sum_{i=1}^{M} r(-i,-j)\left[f(i,j)+v(i,j)\right]\}$$

$$= \sum_{j=1}^{N}\sum_{i=1}^{M} r(-i,-j)f(i,j), \tag{4.7}$$

and

$$\mathrm{Var}\{Z_{rg}(0,0)\} = E\{(\sum_{j=1}^{N}\sum_{i=1}^{M} r(-i,-j)g(i,j) - E\Big[\sum_{j=1}^{N}\sum_{i=1}^{M} r(-i,-j)g(i,j)\Big])$$

$$\cdot(\sum_{l=1}^{N}\sum_{k=1}^{M} r(-k,-l)g(k,l) - E\Big[\sum_{l=1}^{N}\sum_{k=1}^{M} r(-k,-l)g(k,l)\Big])\}$$

$$= E\{\Big[\sum_{j=1}^{N}\sum_{i=1}^{M} r(-i,-j)v(i,j)\Big]\Big[\sum_{l=1}^{N}\sum_{k=1}^{M} r(-k,-l)v(k,l)\Big]\}. \tag{4.8}$$

We define the second-order statistics of the noise by

$$Q(i,j;k,l) = E\{v(i,j)v(k,l)\}.$$

Eq.(4.8) becomes

$$\mathrm{Var}\{Z_{rg}(0,0)\} = \sum_{i,j}\sum_{k,l} r(-i,-j)r(-k,-l)Q(i,j;k,l). \tag{4.9}$$

Substituting Eqs.(4.7) and (4.9) into Eq.(4.3) yields

$$SNR(0,0) = \frac{\left[\sum_{i,j} r(-i,-j)f(i,j)\right]^2}{\sum_{i,j}\sum_{k,l} r(-i,-j)r(-k,-l)Q(i,j;k,l)}. \qquad (4.10)$$

It is convenient to represent the image by a vector f. Correspondingly, r(.,.) can be written as r, and Q(i,j;k,l) as Q. Then the optimal filter problem reduces to the problem of finding r such that the SNR in Eq.(4.11) is maximized given f and Q,

$$SNR(0,0) = \frac{(r^T f)^2}{r^T Q r}. \qquad (4.11)$$

Let

$$\frac{\partial SNR(0,0)}{\partial r} = 0.$$

We obtain

$$f^T r f r^T Q r = f^T r r^T f Q r. \qquad (4.12)$$

Observing that $f^T r$ is a scalar and $f^T r \neq 0$, we have

$$f r^T Q r = r^T f Q r. \qquad (4.13)$$

After some simple rearrangements, Eq.(4.13) becomes

$$Q r = \frac{r^T Q r}{r^T f} f. \qquad (4.14)$$

Define

$$\frac{\mathbf{r}^T\mathbf{Q}\mathbf{r}}{\mathbf{r}^T\mathbf{f}} \triangleq \alpha.$$

Solving for r in Eq.(4.14) yields

$$\mathbf{r} = \alpha\mathbf{Q}^{-1}\mathbf{f}. \tag{4.15}$$

Substituting Eq.(4.15) into Eq.(4.11), we get

$$\max_{\mathbf{r}} \text{SNR}(0,0) = \mathbf{f}^T\mathbf{Q}^{-1}\mathbf{f}. \tag{4.16}$$

Equation (4.16) produces the best-case detection insofar as it maximizes the SNR at the position where the signal would be centred in the image.

It is interesting to note that if v is stationary white noise then $\mathbf{Q}=\sigma_v^2\mathbf{I}$, (where I is an NMxNM identity matrix for images of NxM in dimension, and σ_v^2 is the variance of v). Then the optimal filter reduces to

$$\mathbf{r} = (\alpha/\sigma_v^2)\mathbf{f}. \tag{4.17}$$

If we define $\mathbf{r}=(\alpha/\sigma_v^2)\mathbf{r}_1$, we get

$$\mathbf{r}_1 = \mathbf{f}. \tag{4.18}$$

This is also an optimal filter in the sense that it maximizes the SNR. Note that, from Eq.(4.11), the value of SNR is not affected by scaling r. This implies that we may simply let r_1=r for stationary white noise, that is

$$r = f, \tag{4.19}$$

or

$$r(-i,-j) = f(i,j). \tag{4.20}$$

This result is the well known "matched filter theorem" (Rosenfeld and Kak, 1982) which states that the "best" filter given SNR is the image itself if the noise is a stationary white process. Due to its simplicity and the ease of implementation, matched filters are widely used even if the noise is nonwhite. The use of matched filters in nonwhite noise cases will obviously yield nonoptimal filtering results for the SNR.

Vijaya Kumar and Carroll (1984) derived the "worst-case" detection based on the matched filter, and they defined the following factor for quantitative study of the loss of optimality when non-optimal filters are used,

$$S = \frac{(f^T f)^2}{(f^T Q f)(f^T Q^{-1} f)}, \tag{4.21}$$

and also show, via the Kantorovich inequality, that

$$\frac{4}{(\sqrt{\lambda_{min}/\lambda_{max}}+\sqrt{\lambda_{max}/\lambda_{min}})^2} \leq S \leq 1 \quad , \tag{4.22}$$

where λ_{max} and λ_{min} are the maximum and minimum eigenvalues for the noise covariance (noise autocorrelation) matrix.

The problems with the optimal filter are that it requires: (a) the perfect knowledge of the noise which is normally unavailable; and (b) the inversion of the noise covariance matrix. The latter operation is highly time consuming and unstable for large images, except for some special cases (e.g. when the noise is stationary (Rino, 1970; Gray, 1971)). Consequently our aim in the following sections is to implement an alternative scheme which bypasses these problems, yet still produces better results.

4.3 The Signal Reference Adaptive Postfiltering Technique

We assume an observed image model defined by Eq.(4.1) where v(x,y) is not necessarily white noise. Further to this, we do not care whether v(x,y) is stationary or not. We first consider the cross-correlation of f(x,y) with both sides of Eq.(4.1) to give

$$C_{fg}(x,y) = C_{ff}(x,y) + C_{fv}(x,y) \quad , \tag{4.23}$$

where

$$C_{ab}(x,y) = \iint_A a(u,v)b(x+u,y+v)dudv, \tag{4.24}$$

where A is the domain of integration.

For nonwhite noise, it must be true that

$$\frac{|E[\ C_{fg}(x,y)\ ]|^2}{\mathrm{Var}[\ C_{fg}(x,y)\ ]} \leq \frac{|E[\ Z_{rg}(x,y)\ ]|^2}{\mathrm{Var}[\ Z_{rg}(x,y)\ ]}, \tag{4.25}$$

where the subscript r denotes the optimal filter. The equality is achieved only when v(x,y) is white noise or f happens to be one of the eigenvectors of the noise covariance matrix (see Eq.(4.21)).

It is clear from Eqs.(4.23) and (4.25) that, ideally, if v=0, the SNR would be infinite, and C_{fg} represents a perfect match. Our objective is therefore to design a postfilter h(Θ) which filters C_{fg} to result in an image as close as possible to the signal autocorrelation image C_{ff}. We shall attempt to suppress the error introduced by v(x,y) in the hope of increasing the SNR to a higher value than achieved with a simple cross-correlator. We do this by minimizing the function J(Θ) with respect to the parameter vector Θ,

$$J(\Theta) = E\{[\ C_{tt}(x,y) - \hat{C}_{tg}(x,y)\ ]^2\}, \tag{4.26}$$

where

$$\hat{C}_{tg}(x,y) \triangleq h(\Theta) * C_{tg}(x,y)\ . \tag{4.27}$$

Before proceeding further, some discussion of the signal reference adaptive identification procedure is in order. As shown in Figure 4.1, this procedure consists of two parts, namely, the training mode and the the operation mode. In the training mode, we embed a reference (or standard) image in the noisy background. The parameters of $h(\Theta)$ are then "tuned" by an identification algorithm based on $J(\Theta)$. After the parameters have been decided the system switches to the operation mode. By the nature of $J(\Theta)$, the postfilter $h(\Theta)$ so designed (i.e. h minimizes $J(\Theta)$ when the reference image is present) will reduce the error due to $v(x,y)$. Therefore, in a binary detection problem, if only noise is present, we have only $C_{tv}(x,y)$, and the estimate $\hat{C}_{tg}$ is, in fact, $h*C_{tv}$ which will not minimize $J(\Theta)$.

Since in the case of the detection of an isolated signal in an isolated image, we have no random series (i.e. the absence of the respective probability density functions), we can work only with spatial averages. That is, for digitized images, the statistics are determined from the

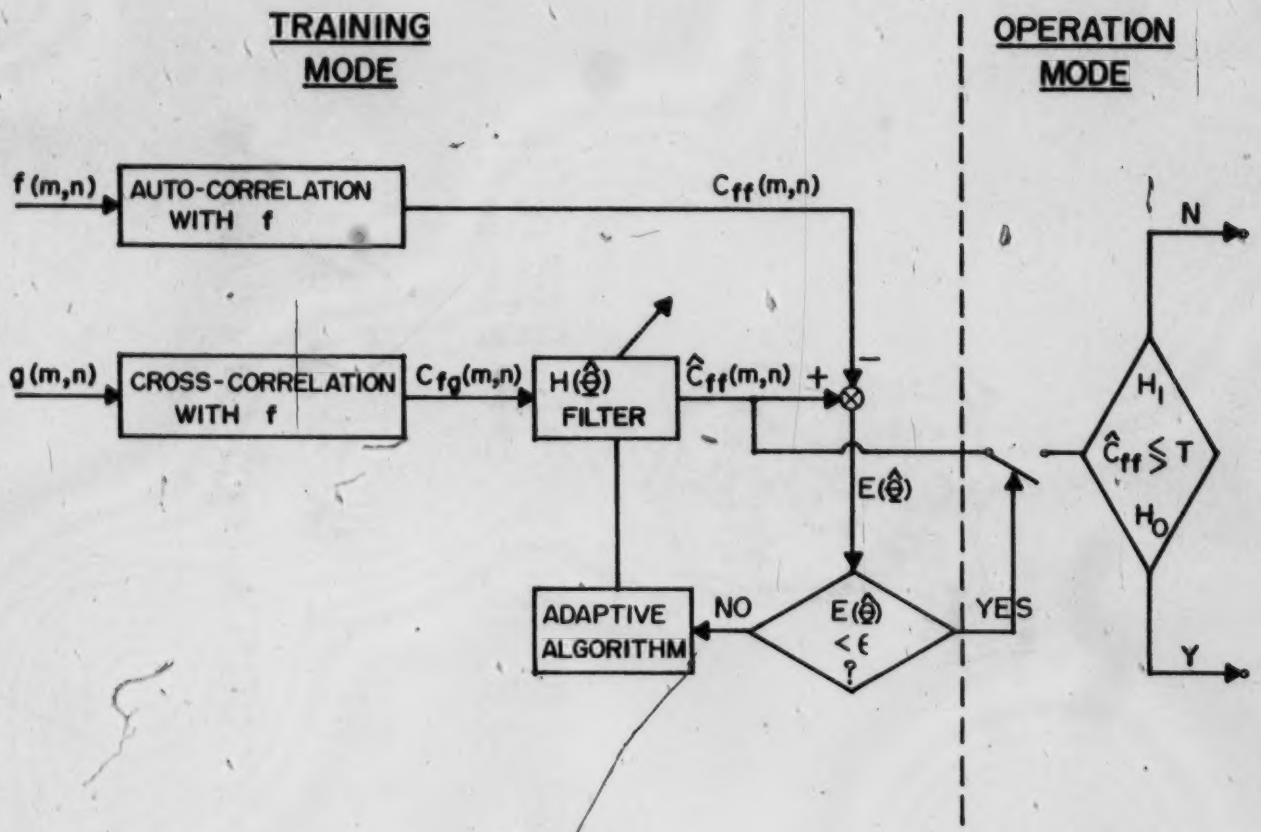


Figure 4.1 The reference adaptive postfiltering matching technique

neighbourhood of each pixel. This reduces Eq.(4.26) to

$$J(\Theta)=\frac{1}{MN}\Big[\sum_{m}\sum_{n}\big(\ C_{ff}(m,n)\ -\ \hat{C}_{fg}(m,n)\ \big)^2\Big], \tag{4.28}$$

where M,N are the image dimensions.

In the digital signal context, the convolution of Eq.(4.27) can be arranged as vector multiplication,

$$\hat{C}_{fg}(m,n)\ =\mathbf{H}^T(\Theta)\mathbf{C}_{fg}(m,n). \tag{4.29}$$

Differentiating Eq.(4.28) with respect to the parameter vector Θ yields

$$\frac{\partial J(\Theta)}{\partial \Theta}=\frac{-2}{MN}\sum_{m}\sum_{n}\frac{\partial H(\Theta)}{\partial \Theta}\mathbf{C}_{fg}(m,n)\Big[\ C_{ff}(m,n)\ -\ \mathbf{H}^T(\Theta)\mathbf{C}_{fg}(m,n)\Big], \tag{4.30}$$

where $H(\Theta)$ is the vector form of $h(\Theta)$. However, since $J(\Theta)$ is a non-linear function of Θ we use non-linear programming techniques for optimization of Eqs.(4.28) and (4.30). For simplicity we set $H(\Theta)=\Theta$ so that the estimate of the autocorrelation function becomes

$$\hat{C}_{fg}(m,n)\ =\ \Theta^T\mathbf{C}_{fg}(m,n)\ . \tag{4.31}$$

Eq.(4.28) may be rewritten as

$$J(\Theta)\ =\ \frac{1}{MN}\sum_{m}\sum_{n}\Big[\ C_{ff}(m,n)\ -\ \Theta^T\mathbf{C}_{fg}(m,n)\ \Big]. \tag{4.32}$$

Let the derivative Eq.(4.30)

$$\frac{\partial J(\Theta)}{\partial \Theta} = 0, \tag{4.33}$$

which gives

$$\sum_m \sum_n C_{ff}(m,n)\ C_{fg}(m,n) = \sum_m \sum_n \Theta^T C_{fg}(m,n)\ C_{fg}(m,n). \tag{4.34}$$

Note that $\Theta^T C_{fg} = C_{fg}^T \Theta$ which is a scalar. Rearranging Eq.(4.34) yields

$$\sum_m \sum_n C_{ff}(m,n)\ C_{fg}(m,n) = \left[\sum_m \sum_n C_{fg}(m,n) C_{fg}^T(m,n)\right]\Theta. \tag{4.35}$$

The vector Θ minimizing $J(\Theta)$ is thus

$$\hat{\Theta} = \left[\sum_m \sum_n C_{fg}(m,n)\ C_{fg}^T(m,n)\right]^{-1}\left[\sum_m \sum_n C_{ff}(m,n)\ C_{fg}(m,n)\right], \tag{4.36}$$

where $\wedge$ indicates the value of Θ that minimizes $J(\Theta)$. Clearly the $\hat{\Theta}$ obtained by Eq.(4.36) is valid over the entire image field.

Before we proceed to perform the simulation, in the next section we shall briefly discuss some aspects of background prefiltering and the ROC analysis which will be used in the simulation to evaluate the detector performance.

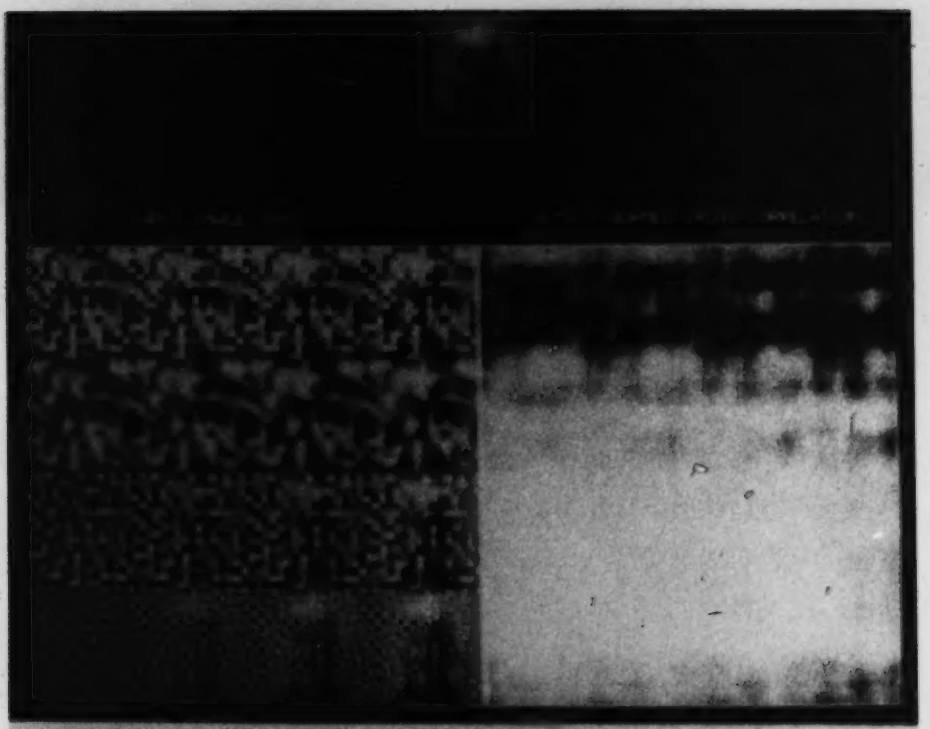
4.4 Detection Criteria: ROC Analysis

The importance of image or background prefiltering on signal detection is illustrated in Figure 4.2a. In the top row of Figure 4.2a(left) a face signal is embedded in bandlimited noise having different amplitudes. The following rows correspond to low-, middle- and high-pass Gaussian filter versions of the image noise-plus-signal, to result in different levels of signal detectability as a function of the band-pass range. Define

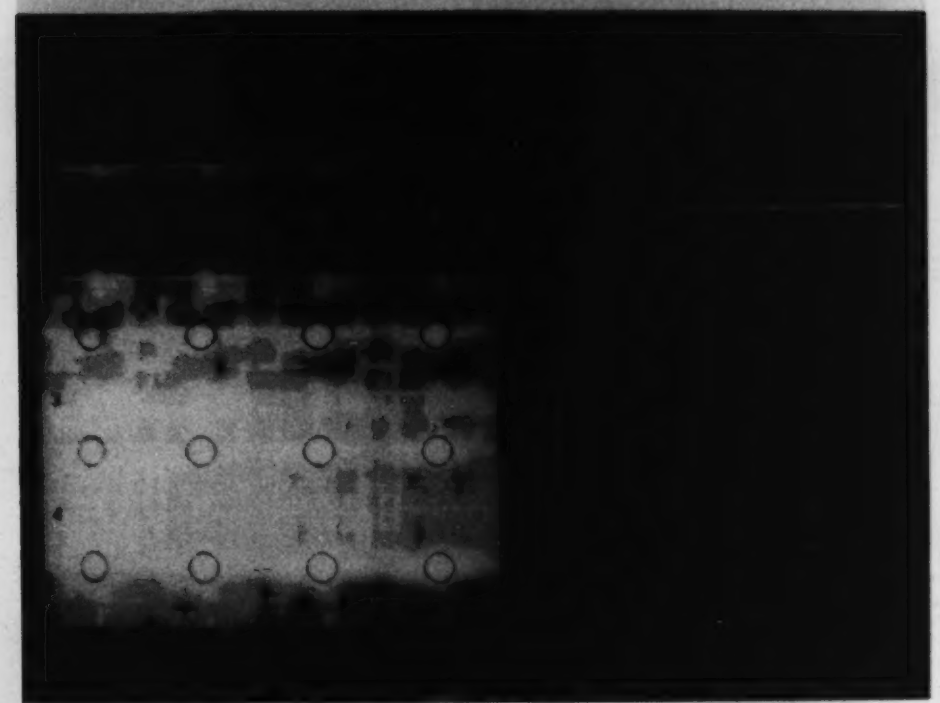
$$g_a(d) = \exp[-\alpha^2(d-d_0)^2], \tag{4.37}$$

where $d \triangleq (u,v)$, $d_0 \triangleq (u_0,v_0)$, and (u,v) are the two-dimensional spatial frequency spectrum. This Gaussian filter enhances detection in terms of the degree to which it filters out frequency components of the noise not common to the signal. In the case of the face image, which is prominently low-to-middle pass in its predominant energy components, best detection occurs in the background by a low-pass filter—as shown in Figure 4.2a: left bottom row. That is, normalized cross-correlations were calculated over each row of the images shown in Figure 4.2a *independently*. The detection efficiency is therefore determined not by the absolute grey level intensity across the rows, but the relative intensities within each row, or a set of four images having different mixing coefficients.

Figure 4.2 Signal, its noise corrupted versions, the cross-correlations, and the ROC curves.
(A) A 64x64 8-bit grey level face image (top: signal) embedded in four different backgrounds (left). Left: First row corresponds to the checkerboard noise, while row 2 to 4 are low-, middle-, and high-pass isotropic Gaussian filtered versions of the noisy images. Right: Results of the normalized cross-correlation between the face image (top of this figure) and the associated face-plus-noise images (left). The pixel intensity is directly proportional to the response of the matched filter,which, in turn, indicates the likelihood of the signal's presence.
(B) Four different background noise types (rows of (A), bottom) are cross-correlated with the signal and the signal region is defined by a circle of 8-pixel radius (depending on applications) about the its centre in the cross-correlation images (left). This set of images are called likelihood maps. The resultant detectability scores are shown to the right by the receiver operating characteristic (ROC) curves for the likelihood maps (left).



A



B

To define the different levels of signal detectability, or the "detection efficiency" of the filter, we have employed the following criteria. In these cases the image g(x,y) was composed by:

$$g(x,y) = \alpha[f(x,y)-E\{f(x,y)\}]+v(x,y), \qquad (4.38)$$

where f(x,y) and v(x,y) are the signal and noise (we use the random checkerboard signal which is nonwhite and nonstationary[5] as the noise in our simulations) respectively, $E\{f\}$ corresponds to the (ensemble) average, or, DC of the signal. Here, going across columns in Figure 4.2a α = 0, .33, .67 and 1.0 respectively. We have computed the normalized cross-correlation images according to (A2.1)-(A2.3)and the results are shown in Figure 4.2a (on the right).

For evaluation we define the signal area in terms of a circular area of radius d about the specified signal centre d_0. The degree to which the filter detects the signal within this region ("hit") is defined, with respect to a threshold λ, by

[5]At the level of pixel resolution, such noise contains correlated information: within a given checker micro-pattern area all pixels assume the same intensity. This introduces non-zero elements in the corresponding checkerboard covariance matrix indicating the non-stationarity nature of the noise. It should also be noted that such noise does not have a flat power spectrum.

$$p[f/(f+v)] \triangleq \frac{\{Z(x,y)>\lambda\}}{A_s(x,y)}\Big|_{(x,y)\in A_s(x,y)} = p(Z(x,y)>\lambda)\Big|_{(x,y)\in A_s(x,y)}, \tag{4.39}$$

where $A_s(x,y)$ defines the signal area. Similarly, the "false alarm" rate is defined by

$$p[f/v] \triangleq \frac{\{Z(x,y)>\lambda\}}{\bar{A}_s(x,y)}\Big|_{(x,y)\in A_s(x,y)} = p(Z(x,y)>\lambda)\Big|_{(x,y)\in A_s(x,y)}, \tag{4.40}$$

where $\bar{A}_s(x,y)$ is the complement $A_s(x,y)$, and $Z(x,y)$ corresponds to the output of a given convolution or filtering process, as in Eq.(4.2), etc..

By varying λ over a range of the 8-bit intensity values we have produced ROC (receiver-operating-characteristics), as shown in Figure 4.2b (on the right) with the associated threshold values in Figure 4.2b (on the left). The area under these ROC curves, namely,

$$P_A = \int_0^1 P[f/(f+v)]dp(f/v), \tag{4.41}$$

determines a measure of the detectability of the signal-independent of the criterion value chosen. This area varies between 0-1.0, usually 0.5 to 1.0, where 1.0 implies perfect detection efficiency. Further, this measure permits direct comparison between the different prefilters, and, to the unfiltered or direct cross-correlation case.

Though we illustrate, here, the benefit of prefiltering

in improving detection efficiency with cross-correlation, we have not generated an objective technique for generating the "ideal postfilter".

4.5 Simulations

In this section we shall investigate the effectiveness of the proposed adaptive matched filtering technique using computer simulation. without loss of generality we choose to have seven elements over 5x5 symmetrically structured kernal (see Figure 4.3). The cross correlation vector C_{fg}(see Eq.(4.29)) in this configuration is

$$C_{fg}(m,n)=\begin{bmatrix} C_{fg}(m-2,n-2)+C_{fg}(m+2,n-2)+C_{fg}(m+2,n+2)+C_{fg}(m-2,n+2) \\ C_{fg}(m-1,n-2)+C_{fg}(m+1,n+2)+C_{fg}(m+1,n-2)+C_{fg}(m-1,n+2) \\ C_{fg}(m,n-2)+C_{fg}(m+2,n)+C_{fg}(m,n+2)+C_{fg}(m-2,n) \\ C_{fg}(m-2,n-1)+C_{fg}(m+2,n-1)+C_{fg}(m+2,n+1)+C_{fg}(m-2,n+1) \\ C_{fg}(m-1,n-1)+C_{fg}(m+1,n-1)+C_{fg}(m+1,n+1)+C_{fg}(m-1,n+1) \\ C_{fg}(m,n-1)+C_{fg}(m+1,n)+C_{fg}(m,n+1)+C_{fg}(m-1,n) \\ C_{fg}(m,n) \end{bmatrix}$$

As shown in Figure 4.3, we have employed a specific mask of Θ defined by

$$\Theta = [\theta_1 \;\; \theta_2 \;\; \theta_3 \;\; \theta_4 \;\; \theta_5 \;\; \theta_6 \;\; \theta_7]^T. \qquad (4.42)$$

In this example the parameter vector has the following value

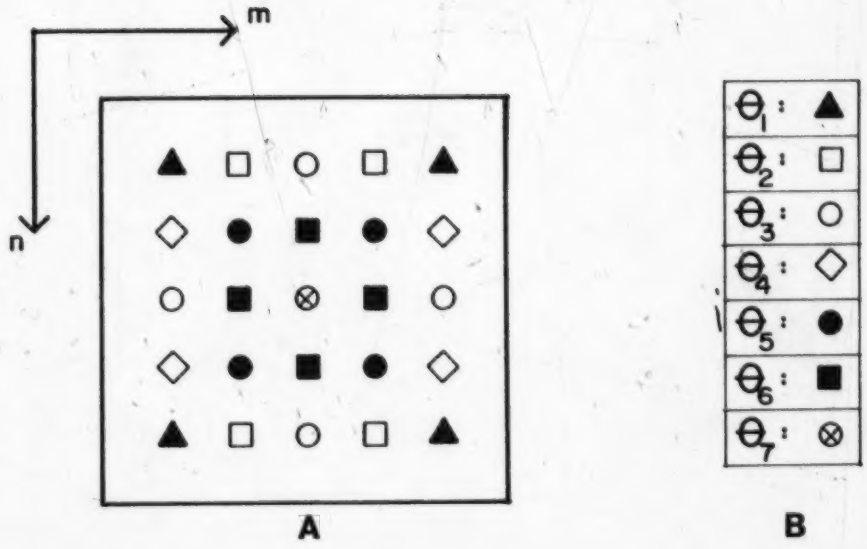


Figure 4.3 The postfilter parameter configuration in a 5x5 window

$$\hat{\theta} = [0.077, 0.0217, -0.164, 0.175, 0.0167, -0.295, 1.352]^T, \quad (4.43)$$

with respect to the filter profile shown in Figure 4.3a which depicts the θ values over the 5x5 kernel structure used.

Consistent with the idea that $h(\theta)$ should be estimated adaptively, as shown in Figure 4.1, we employed a nonlinear programming algorithm to find the local minimum of Eq.(4.28) and so to improve our estimates of θ. This was based on the finite difference Levenberg-Marquardt routine for solving non-linear least squares problems. A sequence of approximations to the minimum point is generated by

$$\hat{\theta}_{i+1} = \hat{\theta}_i - [\alpha_i D_i + K_i^T K_i]^{-1} K^T{}_i J(\hat{\theta}_i), \quad (4.44)$$

where K_i is the numerical Jacobian matrix evaluated at θ_i; D_i is a diagonal matrix; α_i is a positive scaling constant. This results in the estimate of θ:

$$\hat{\theta} = [0.087, 0.023, -0.164, 0.174, 0.0171, -0.293, 1.343]^T, \quad (4.45)$$

which differs only slightly from Eq.(4.43).

Figure 4.1 is intended for more general adaptive estimation of $H(\theta)$ than the particular example described above. In fact the procedure is envisaged to be divided into

two parts: (a) the training mode and, (b) the normal detection operator. In training mode a "standard" image f and its autocorrelation C_{ff}, and the observed image g and its cross-correlation C_{fg} (with f) are given. The adaptive algorithm uses $E(\Theta)$ (i.e. $J(\Theta)$ here) as the input to adjust the parameters of the postfilter until $E(\Theta)$ is within the error limit given as a constant ϵ. Upon the parameter vector Θ is determined, the system is automatically oriented to normal operation mode for image matching.

Figure 4.4 shows the ROC curves (and detectability levels) associated with each of the matching techniques for the detection of the face image embedded in noise (random checkerboard signal) corresponding to row 4, column 1 of Figure 4.2b(left). As can be seen from these results, the postfilter derived here is 91% as efficient as the autocorrelation result. However, they compare favourably to the normalized cross-correlation and, in particular, the direct cross-correlation values, $C_{ff} = 0.98$, $C_{fg1} = 0.89$, $C_{fg2} = 0.88$, $C_{fgn} = 0.80$, and $C_{fg} = 0.72$. Here, C_{ff} indicates signal autocorrelation; C_{fg1} and C_{fg2} are the estimated cross-correlation values using the parameter given by Eqs.(4.43) and (4.45) respectively; C_{fg} is the direct cross-correlation of signals f and g, and C_{fgn} indicates the normalized value (see APPENDIX II for the definition) of C_{fg}.

Since in matched filtering the shape of the

cross-correlation image proves to be of utmost importance, we have further compared these results by correlating the autocorrelation image with the others via the standardized Pearson's correlation coefficient defined by

$$-1 \leq C(I_1, I_2) = \frac{\mathrm{cov}(I_1, I_2)}{\sigma_{I1}\sigma_{I2}} \leq +1. \qquad (4.46)$$

This resulted in values of $C(C_{ff}, C_{fg1}) = 0.81$, $C(C_{ff}, C_{fg2}) = 0.81$, $C(C_{ff}, C_{fgn}) = 0.69$ and $C(C_{ff}, C_{fg}) = 0.08$.

It is worth noting that this technique estimates a linear filter $h(\Theta)$ which attempts to minimize the difference between a given signal's auto-correlation function and the associated signal-to-image cross-correlation function in a general setting. This approach considerably improves the direct matched filtering method, and does not require the knowledge of noise statistics which in many cases makes the optimal detection impractical.

Figure 4.4 Simulation results.
On the left, five comparison cross-correlation functions involving the signal (top of Figure 4.2a) and the fourth image in row 1 of Figure 4.2a. Right hand side shows the associated ROC curves and detectability values.

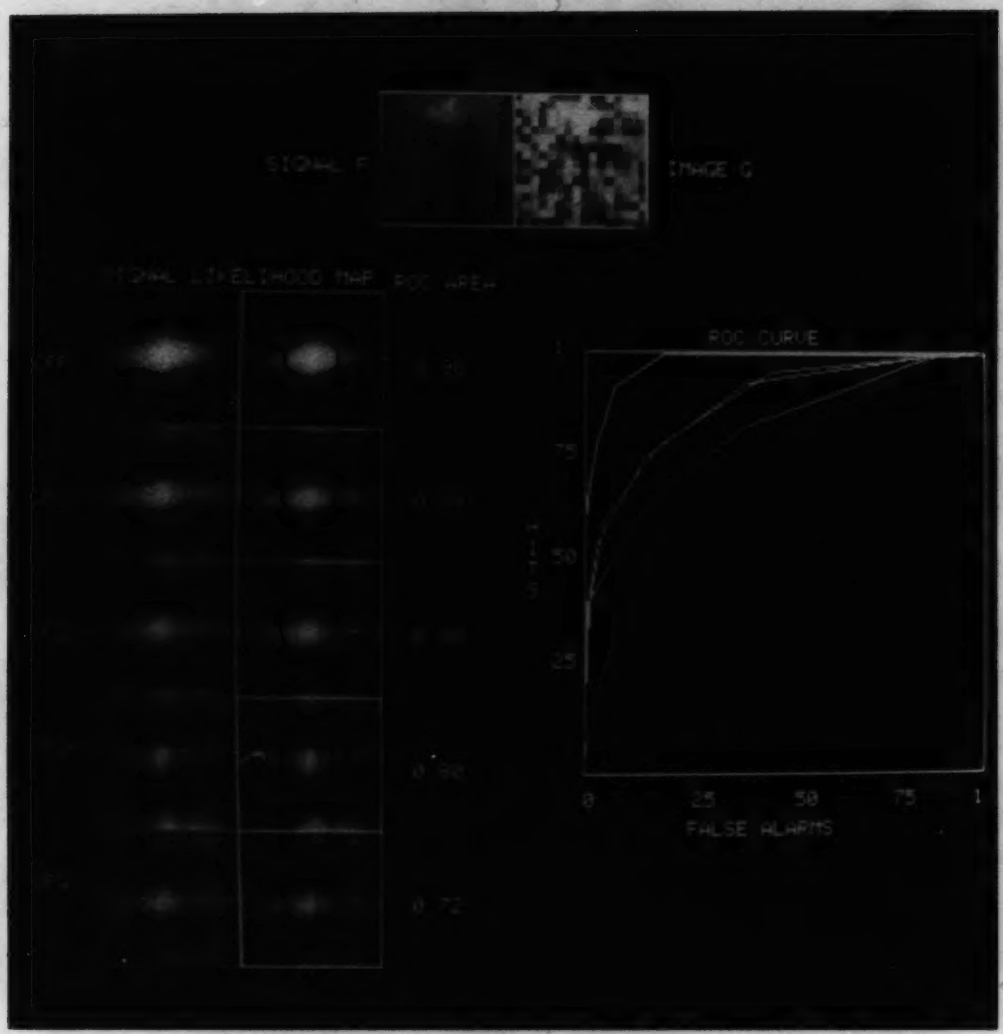


IMAGE G
ROC CURVE
0
25
50
75
1
FALSE ALARMS

CHAPTER V

A HIERARCHICAL APPROACH TO MULTI-OBJECT PATTERN RECOGNITION AND DETECTION IN NOISY BACKGROUNDS

5.1 Introduction

In previous chapters we discussed image restoration and detection in noisy backgrounds. Algorithms were developed to yield images with better visual quality and to detect objects with higher accuracy. It was pointed out that the theory and methods of matched filters have been widely used in signal detection over the past few decades specifically because matched filtering is optimal for the detection of signals corrupted by stationary white noise. Essentially, the matched filtering involves cross-correlation of the known signal with the observed image. The location of the signal, if any, in the image is indicated by the peak value of the cross-correlation. Recently, approaches to reducing the computation time for such matching techniques (Rosenfeld and Kak, 1982) have been developed, including sequential methods (Wong et al., 1978; Munteanu, 1981) and hierarchical approaches (Rosenfeld et al., 1977; Tanimoto, 1981).

The basic principle of the sequential methods is to accumulate the degree of mismatch between the template and each position on the image. Since mismatches usually grow faster than match, the poorly matched area can be detected quickly and rejected at an early stage of the operation. In

this way, the cost of determining a match is cut down considerably.

In the hierarchical approaches, resolution of images and template is reduced by averaging or pyramids. The matching is performed first over the reduced resolution images. If a promising area is detected for a given threshold then full matching is performed for the exact location; otherwise the operation aborts that area and searches for another area. This approach speeds up the matching procedure to some extent. An expected problem with this procedure is that the sensitivity of the procedure is low, which means more false alarms.

Most recently, in searching for a fast detection scheme for robot vision Caelli and Nagendran (1986) have further developed the "edge only matching technique". In this approach, the object being detected is first reduced to an edge only signal. This significantly reduces the redundant information of the signal so that the matched filtering can be performed by a simple array processor at a very high speed. The basic principle, which led to this approach is the perceptual importance of edges in human matching tasks (Caelli and Moraglia, 1986), or, in general, the perceptual salience of image bandpass information for pattern recognition. However, such pre-filtering methods still employ the cross-correlation of the signal with the input image and become quite inefficient when multi-object

detection is required.

Pattern recognition, whether dealing with single or multiple signals, is essentially concerned with the development of classification procedures based upon explicit features of the signal(s) and images. Figure 5.1 illustrates this fundamental operating principle. Usually, the characteristic features of the image are extracted by using some transformation techniques such as orthogonal transformations, statistical transformations, or syntactic transformations. The resulting features may turn out to be, depending on the kind of transformation techniques used, the set of largest eigenvalues (Ahmed and Rao, 1975), or the grey-scale histogram, or the frequency spectrum, or the auto-correlation (Devijer and Kittler, 1982), or syntactic descriptions (Fu, 1982). In actual applications, a set of sample images is used to generate the sample (or reference) features, which make up the feature space.

Since, in most realistic situations, images are corrupted by environmental noise, and the statistical information of the noise is either unknown, or only partially known à priori, it is not feasible to use different sets of noisy samples to train the classifier. Therefore, the training samples are, normally, a set of ideal images free of noise degradation but sometimes involving subjectively chosen geometrical distortions. Trained by such samples, the classifier will produce

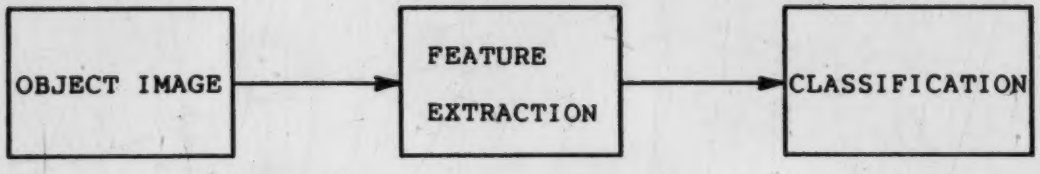


Figure 5.1 The pattern recognition problem

misclassification in a noisy environment. Further to this, ~~in~~ final detection, the performance of the matched filter will also deteriorate if the noise is nonwhite. In such circumstances the perfofmance of the hierarchical approach will become unacceptably poor. One solution is to resort to image restoration techniques for the preprocessing of the noise images, if sufficient information is available.

In the following sections we introduce a linear least squares mapping technique (LLSMT) for classification and demonstrate how noise may confuse the classifier, and lead to misclassification or no decision at all. In pattern classification the visual quality of the restored image is not essential as long as the basic features are retained. To improve this situation we introduce a very simple and fast prefiltering scheme as the first stage in the proposed new hierarchical approach.

5.2 The Linear Least Squares Mapping Technique

In developing discriminant functions based on minimum distance criteria, it is often assumed that the pattern values cluster around their respective mean in feature space. This assumption may not be reasonable in many applications (Ahmed, et al., 1975). In such cases, it is necessary to transform the observed patterns or features belonging to a given class, say, C_i, into a decision space, such that these features are forced to cluster around a

point v_i in the decision space. The tranformation **A** which maps the feature space to the decision space is determined for all the classes involved and is based on a linear least squares error criterion for a given set of unit vectors V_i. These vectors $\{V_i : i=1,2,\ldots,k\}$ span the decision space. The classification is determined by the proximity of the mapped points to the class points in the decision space. For instance, if among all the mapped points in the decision space three points are within a tolerance range around the point v_j then we say that the three points belong to class C_j.

The advantages of using the LLSMT are that

(a) Given a feature vector or an image statistically belonging to class C_i, its distance from v_i will be minimized.

(b) Although the mapped values will change from sample to sample, they will be closer to one another compared to those not in the same class.

(c) Since the decision space is spanned by the vectors V_j the jth row of **A** maps the feature vectors of the jth class to the vector corresponding V_j.

We shall now consider some sets of sample feature vectors $\{f_{kl}^{(s)};\ k=1,2,\ldots,K;\ l=1,2,\ldots,L\}$, where superscript (s) denotes the samples, K is the number of classes and L the number of samples in each class. We seek a transformation **A** which maps $\{f_{kl}^{(s)}\}$ into the unit vectors V_k.

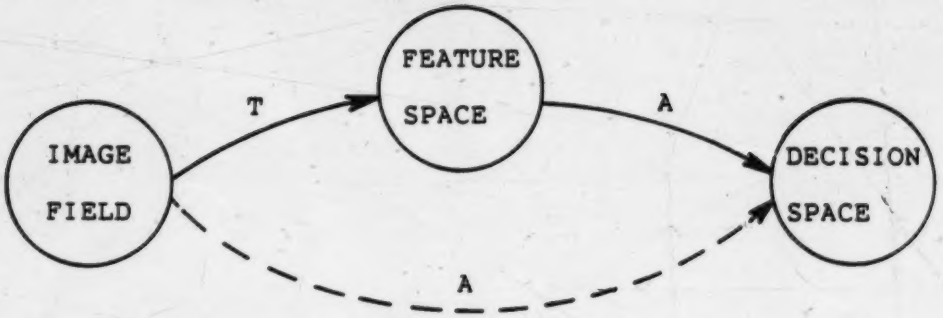


Figure 5.2 Transformation of patterns into decision space

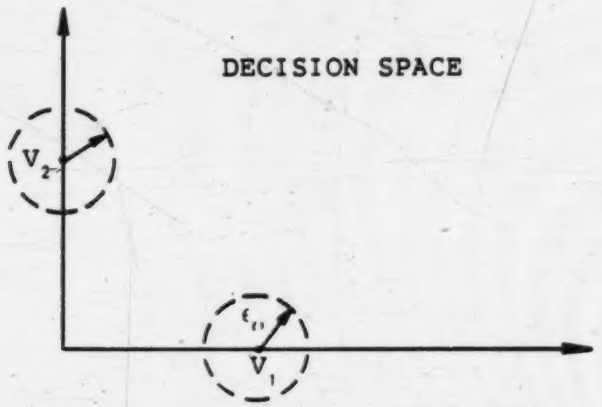


Figure 5.3 Patterns of class 1 centred around the point V_1 in the decision space

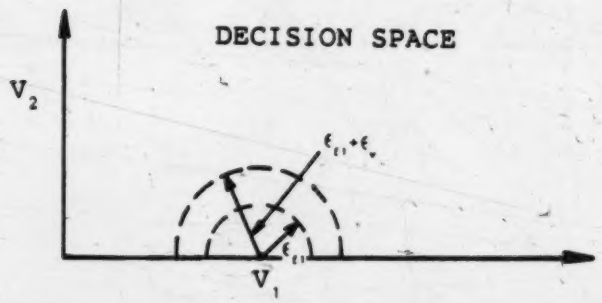


Figure 5.4 Mapping error caused by noise

We define

$$\epsilon_{kl} = \mathbf{A}\mathbf{f}_{kl}^{(s)} - \mathbf{V}_k = \mathbf{L}_{kl} - \mathbf{V}_k. \tag{5.1}$$

The mean-square error of one class, given the samples, is

$$\epsilon_k = \frac{1}{L}\sum_{l=1}^{L} \| \epsilon_{kl} \|^2. \tag{5.2}$$

Substitution of Eq.(5.1) into Eq.(5.2) yields

$$\epsilon_k = \frac{1}{L}\sum_{l=1}^{L}\left[\mathbf{L}_{kl}-\mathbf{V}_k\right]^T\left[\mathbf{L}_{kl}-\mathbf{V}_k\right]$$

$$= \frac{1}{L}\sum_{l=1}^{L}\left[\mathbf{f}_{kl}^{(s)T}\mathbf{A}^T\mathbf{A}\mathbf{f}_{kl}^{(s)} - 2\mathbf{f}_{kl}^{(s)T}\mathbf{A}^T\mathbf{V}_k + \| \mathbf{V}_k \|^2\right]. \tag{5.3}$$

$\mathbf{A}$ must be chosen such that ϵ_k is minimized. This can be obtained by solving the following equation

$$\frac{\partial \epsilon_k}{\partial \mathbf{A}} = 0, \tag{5.4}$$

which is

$$\frac{1}{L}\sum_{l=1}^{L}\frac{\partial\left[\mathbf{f}_{kl}^{(s)T}\mathbf{A}^T\mathbf{A}\mathbf{f}_{kl}^{(s)}\right]}{\partial \mathbf{A}} - \frac{2}{L}\sum_{l=1}^{L}\frac{\partial\left[\mathbf{f}_{kl}^{(s)T}\mathbf{A}^T\mathbf{V}_k\right]}{\partial \mathbf{A}} + \frac{\partial \| \mathbf{V}_k \|^2}{L\partial \mathbf{A}} = 0. \tag{5.5}$$

It is obvious that

$$\frac{\partial\left[\mathbf{f}_{kl}^{(s)T}\mathbf{A}^T\mathbf{A}\mathbf{f}_{kl}^{(s)}\right]}{\partial\mathbf{A}} = 2\mathbf{A}\left[\mathbf{f}_{kl}^{(s)}\mathbf{f}_{kl}^{(s)T}\right], \tag{5.6}$$

and

$$\frac{\partial\left[\mathbf{f}_{kl}^{(s)T}\mathbf{A}^T\mathbf{V}_k\right]}{\partial\mathbf{A}} = \mathbf{V}_k\mathbf{f}_{kl}^{(s)T}, \tag{5.7}$$

and since $\mathbf{V}_k$ is independent of $\mathbf{A}$,

$$\frac{\partial\|\mathbf{V}_k\|^2}{\partial\mathbf{A}} = 0. \tag{5.8}$$

Upon substituting these three identities into Eq.(5.5), we have

$$\mathbf{A}\left[\frac{2}{L}\sum_{l=1}^{L_k}\mathbf{f}_{kl}^{(s)}\mathbf{f}_{kl}^{(s)T}\right] = \frac{2}{L}\sum_{l=1}^{L_k}\mathbf{V}_k\mathbf{f}_{kl}^{(s)T}, \tag{5.9}$$

or

$$\mathbf{A}\left[\sum_{l=1}^{L_k}\mathbf{f}_{kl}^{(s)}\mathbf{f}_{kl}^{(s)T}\right] = \sum_{l=1}^{L_k}\mathbf{V}_k\mathbf{f}_{kl}^{(s)T}. \tag{5.10}$$

Since we are seeking $\mathbf{A}$ such that the total error ϵ is a

minimum for all the classes, we sum both sides of Eq.(5.10) over all k, to get

$$\mathbf{A}\Big[\sum_{k=1}^{K}\sum_{l=1}^{L_k}\mathbf{f}_{kl}^{(s)}\mathbf{f}_{kl}^{(s)T}\Big] = \sum_{k=1}^{K}\sum_{l=1}^{L_k}\mathbf{V}_k\mathbf{f}_{kl}^{(s)T}. \tag{5.11}$$

Solving for **A** gives

$$\mathbf{A} = \Big[\sum_{k=1}^{K}\sum_{l=1}^{L_k}\mathbf{V}_k\mathbf{f}_{kl}^{(s)T}\Big]\Big[\sum_{k=1}^{K}\sum_{l=1}^{L_k}\mathbf{f}_{kl}^{(s)}\mathbf{f}_{kl}^{(s)T}\Big]^{-1}. \tag{5.12}$$

5.3 Estimation Scheme for the Prefilter

Though the above LLSMT technique is envisaged as the first part of our hierarchical detection scheme, it is of particular relevance to note that the transformation matrix **A** is generated by using a set of pattern samples taken from different classes which are "ideal" images since no stochastic degradation is present. However, it has already been pointed out that images or patterns recorded by media inevitably suffer from various degradations, such as blurring, noise, and geometric distortion. In this study, however, we restrict our attention to additive noise. We rewrite the discrete observed image model as

$$\mathbf{g} = \mathbf{f}+\mathbf{v}, \tag{5.13}$$

where **g**, **f**, and **v** are the vector forms of g(m,n), f(m,n),

and v(m,n) with variance σ_v^2, respectively. We assume that the feature vector is the image **g** itself, which may be considered as the feature vector containing the full information. Premultiplying on both sides of Eq.(5.13) by the matrix **A**, we get

$$\mathbf{Ag} = \mathbf{Af}+\mathbf{Av}. \tag{5.14}$$

Ideally, if **g** were not corrupted by additive noise, we would have

$$\mathbf{Ag} = \mathbf{Af} = \mathbf{V}_k+\epsilon_f, \tag{5.15}$$

where $\mathbf{V}_k$ is an unit vector in the decision space. Since **A** is trained by only a limited set of object image samples, for an arbitrary image vector **f**, an error ϵ_f, which is different from the minimum ϵ will be incurred. Consequently, the classified objects, represented by vectors in the decision space, will cluster closely about $\mathbf{V}_k$ given ϵ_f (see Figure 5.3 for illustration) as long as the statistics of the object image **f** are similar to that of the training set.

Given an adequate feature list (vectors) for input patterns, it is well established that, when the object images are not corrupted, the LLSMT is useful and effective. However, in contrast to such ideal cases, we obtain, from

Eq.(5.13) and Eq.(5.14)

$$\mathbf{Ag} = \mathbf{V}_i + \epsilon_t + \epsilon_v, \tag{5.16}$$

where $\epsilon_v = \mathbf{Av}$ which will affect the classification as a function of the severity of the noise corruption (see Figure 5.4). Indeed, under these conditions, the use of **A** as determined by Eq.(5.12), may lead to misclassification. We therefore resort to estimation techniques in order to improve the performance of the classifier. This can be done by using prefiltering to initially restore the image.

Unlike the usual situation in image restoration, in the present context, the main concern is to develop prefiltering techniques to suppress noise and in order to singularly improve classification performance without necessarily improving image "quality" per se.

For these reasons we have modified the Wallis' (1976) restoration algorithm. This is a simple algorithm based on knowledge of local mean and variance, and a desired local variance $\sigma_d^2(m,n)$. The estimated image pixel at (m,n) is given by

$$\hat{f}(m,n) = g_m(m,n) + K(m,n)[g(m,n) - g_m(m,n)], \tag{5.17}$$

where

$$K(m,n)=\frac{\sqrt{\sigma_d^2(m,n)}}{\sqrt{\sigma_g^2(m,n)}}\ , \tag{5.18}$$

$$\sigma_g^2(m,n)=\frac{1}{(2i+1)(2j+1)}\sum_{k=m-i}^{i+m}\sum_{l=n-j}^{i+n}\left[g(k,l)-g_m(m,n)\right]^2\ , \tag{5.19}$$

and

$$g_m(m,n)=\frac{1}{(2i+1)(2j+1)}\sum_{k=m-i}^{i+m}\sum_{l=n-j}^{i+n}g(k,l). \tag{5.20}$$

These local statistics are calculated at every pixel over a window area which normally covers 3x3 or 5x5 pixel neighbourhood. In our application, however, we propose the following modified prefiltering algorithm:

$$\hat{f}(m,n)=g_m(m,n)+K[g(m,n)-g_m(m,n)], \tag{5.21}$$

where

$$K=\frac{\sigma_f^2}{\sigma_g^2}\ , \tag{5.22}$$

which is a constant, and

$$\sigma_f^2=\sigma_g^2-\sigma_v^2. \tag{5.23}$$

Further,

$$\sigma_g^2 = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[g(k,1)-\bar{g}(m,n)\right]^2, \tag{5.24}$$

$$\bar{g}(m,n) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}g(k,1). \tag{5.25}$$

In this case, we have assumed the additive noise to be white, with a constant variance of σ_v^2. This scheme has a distinct computational advantage over that of Wallis'. The computations of local σ_d^2 and σ_g^2 are not required. Only the local mean $g_m(m,n)$ is to be computed. In the simulations, we shall see that this algorithm restores images with reasonable quality.

5.4 Implementation of the Hierarchical Detection Technique

As shown in the Figure 5.5, the entire operation is basically divided into three sections:

1. Image prefiltering,
2. Feature encoding,
3. Matched filtering for detection.

We first partition a given image into sub-images each of which may or may not contain an object, and is contaminated by Gaussian white noise. The sub-images are prefitered by using Eq.(5.21). Then the LLSMT is performed over each of the sub-images, and the estimators introduced

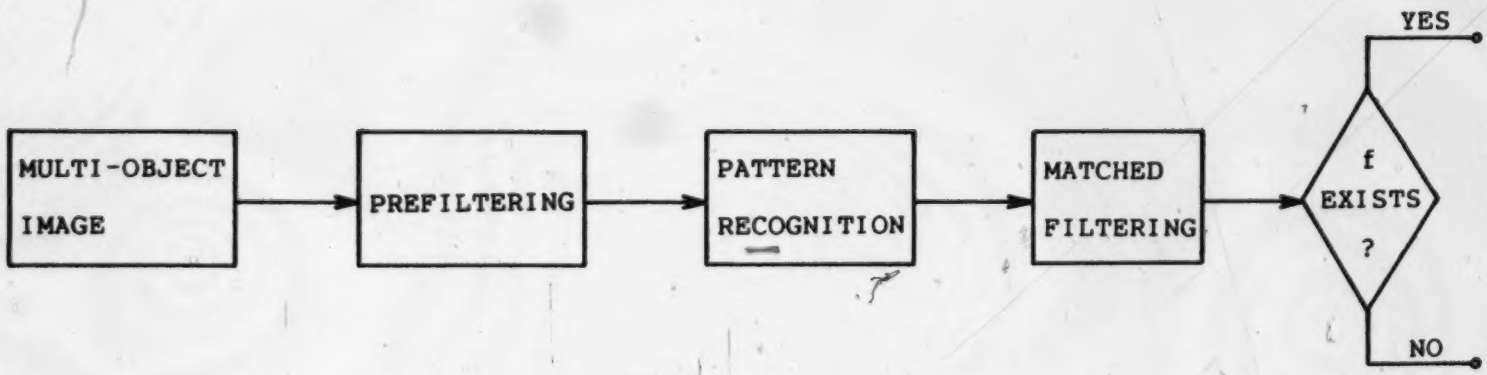


Figure 5.5 The hierarchical detection approach

in the last section are used to reduce the classification error. Having been properly classified, each sub-image is labelled according to the class to which it belongs. In the detection part, we use the matched filter to detect a known object in a class with those sub-images having the same class label.

5.5 Simulations

In order to verify the effectiveness of the proposed hierarchical procedure for detecting multiple objects in an image scene, a computer simulation is carried out according to the order of processes defined in Figure 5.5. The detailed process is depicted by the flow diagram in Figure 5.6.

The problem under consideration is that of three-class multi-object detection. The three classes are "Bird-class", "Face-class", and "Leaf-class". In order to calculate the transformation matrix **A**, we use eight samples for each of the three classes of patterns (Figure 5.7a).

As already pointed out, the direct use of the entire image as the feature space will result in high dimensionality and therefore, in practice, images are usually transformed by different methods to establish the feature space. For instance, an NxN image field f_{ij} if lexicographically represented by an image vector will become a vector f_{ij} of N^2 elements. In contrast, the common feature

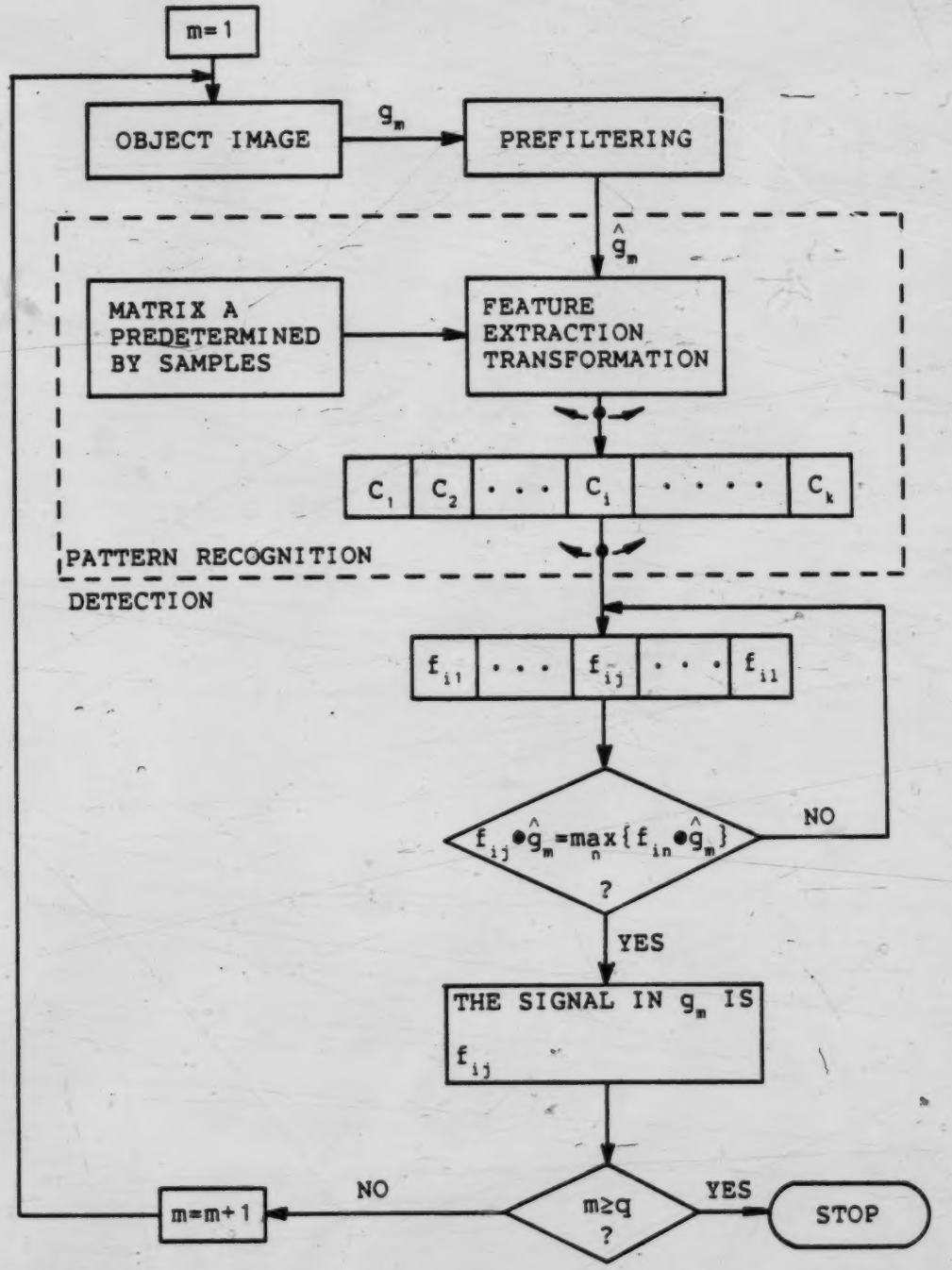


Figure 5.6 The flow diagram of the hierarchical approach

extraction techniques (see Figure 5.1) only use a small set of essential characteristics to represent the original image. The feature vector contains J elements which are considerably less than N^2 of the image vector f_{ij}, thus reducing the computational burden. For the purpose of demonstration in this section, however, we shall simply use the image vector f_{ij} as a feature vector considering a variety of feature represention which are more or less controversial and necessary only when large image sizes and large numbers of samples and classes are involved. Indeed the original input image vector f_{ij} contains the maximum available image information. The operation we choose is the dashed line from image space directly to the decision space shown in Figure 5.2. The simplest orthonormal vectors V_i for the three dimensional decision space are:

$$V_1=\begin{bmatrix}1\\0\\0\end{bmatrix},\ V_2=\begin{bmatrix}0\\1\\0\end{bmatrix},\ V_3=\begin{bmatrix}0\\0\\1\end{bmatrix}.$$

Each sample image in this simulation is of 64x64 in dimension. Thus transformation matrix **A** is of size 3x4096. Looking at Eq.(5.12), the expression for the matrix **A**, we find that in order to compute **A** we first have to compute the inverse matrix $\left[\sum_{k=1}^{K}\sum_{l=1}^{L} f_{kl}^{(s)} f_{kl}^{(s)T}\right]^{-1}$. It seems that the inverse

matrix would be difficult to handle because of its enormous size (4096x4096). However, in this simulation, we have only three classes (K=3) and eight samples (L=8) per class. Consequently, the matrix inversion can be relatively easily solved by using the singular value decomposition (SVD) method (see the APPENDIX III, for details),

We define

$$T^{-1} \triangleq \left[\sum_{k=1}^{K}\sum_{l=1}^{L} f_{kl}^{(s)} f_{kl}^{(s)T}\right]^{-1}, \tag{5.26}$$

by the method introduced in APPENDIX III, we can decompose this as

$$T^{-1} = Q\Phi\Sigma^{-2}\Phi^{T}Q^{T}, \tag{5.27}$$

where Q is a 4096x24 matrix and is given by

$$Q=\left[f_{11}^{(s)}, f_{12}^{(s)}, \ldots, f_{1L}^{(s)}; f_{21}^{(s)}, f_{22}^{(s)}, \ldots, f_{2L}^{(s)}, \ldots; f_{K1}^{(s)}, f_{K2}^{(s)}, \ldots, f_{KL}^{(s)}\right], \tag{5.28}$$

Φ is the matrix composed of the eigenvectors of $Q^{T}Q$, and Σ is a diagnal matrix with the eigenvalues of $Q^{T}Q$ as its diagonal elements. From Eq.(5.27) we see that in order to compute T^{-1}, it is necessary to compute only some matrix products and the inverse square of Σ. The size of Σ is 24x24 in this case. A computer program is generated according to following steps to obtain the matrix A,

(a) Calculate VF, where VF $\triangleq$

$$\left[\sum_{k=1}^{K}\sum_{l=1}^{L} V_k f_{kl}^{(s)T}\right] ;$$

(b) Construct $Q \epsilon R^{4096 \times 24}$;

(c) Compute Σ using the SVD subroutine;

(d) Solve for T^{-1} by using Eq.(5.27);

(e) Solve for matrix **A** by using Eq.(5.12).

In Table 5.1 we have listed the mapping results for the training samples which show that the matrix **A** is able to map each sample, say, $f_{ij}^{(s)}$, remarkably close to the location corresponding to V_i. To show that noise affects the classification, we first add white noise with zero mean and $\sigma_v^2 = 1275$ to the same set of images (see Figure 5.7b), then the matrix **A** is applied to these noisy samples. The results are listed in Table 5.2 which, as expected, clearly show that with the additive white noise the performance of the classifier deteriorates for the given mapping **A**.

A multi-object image field is shown in Figure 5.8a, this image consists of fifteen objects belonging to three classes. In Figure 5.8b the object image field is corrupted by white noise with zero mean and variance $\sigma_v^2 = 1275$. Tables 5.3, and 5.4 contain the mapping results for the noise-free object image (Figure 5.8a) and the noisy object image (Figure 5.8b), respectively. Again we see that the mapping matrix **A** is very effective when the object image is free of noise so long as its statistics are similar to that of the training samples, while for the noise contaminated objects the classifier performance deteriorates.

Figure 5.9a shows the object image restored by using the proposed prefiltering algorithm. Table 5.5 lists the mapping results for the restored object image (Figure 5.9a). These results show considerable improvement on the accuracy of the classfication and they are consistent with the prefiltered object image.

After the objects have been classified, we perform matched filtering technique (normalized cross-correlation) to detect the exact objects present in the image field. For an overview, we list all the 120 cross-correlation peak values in Table 5.6. A close inspection of the table we find that all the diagnal elements have the highest peak correlation values compared to their respective horizontally off-diagnal elements.

To reveal the effectiveness of this approach, we compare the relative detectability by using the ROC introduced in Chapter IV. Since the characteristics of these curves are more or less similar to each other, for demonstration, we show only three typical ROC's together with their likelihood maps (see Figure 5.9b to Figure 5.10b, each corresponds to a class). In each of these figures three pairs of likelihood maps (in this simulation, the images of normalized cross-correlation) are used. Ordered from top down, the first pair are two auto-correlation maps. In the ROC this corresponds to the diagonal straight line showing the same detectability. The second pair consists of the

cross-correlation between the object and the prefiltered object (on the left), and the cross-correlation of the object image with the noisy object (on the right). In the ROC this gives a curve which covers more than 90% of the area indicating that the use of the prefiltering technique results in significant improvement in the detectability. The third pair are the auto-correlation of the object (on the left) and the cross-correlation between the object and the prefiltered object (on the right). The resulting ROC is a curve with small deviations from the straight diagnal line revealing the closeness of the two, which in turn implies that the detectability is improved by using the prefiltered image.

The experiments have shown that using the hierarchical procedure to detect objects in a multiple object image we can save more than 50% computing time compared with the brute force application of the matched filtering technique, and virtually no sacrifice has been made on the overall performance of the detection procedure. For noisy images the prefiltering technique significantly improves the the accuracy of the classifier and reduces false alarms of the detector.

Table 5.1 The mapping results for the training sample images

(a) The Bird-Class Samples

Sample No.	V_1	V_2	V_3
1	1.0062	$6.6X10^{-3}$	$5.2X10^{-3}$
2	1.0047	$4.7X10^{-3}$	$3.7X10^{-3}$
3	1.0030	$3.3X10^{-3}$	$2.5X10^{-3}$
4	1.0052	$5.3X10^{-3}$	$4.2X10^{-3}$
5	1.0056	$5.7X10^{-3}$	$4.6X10^{-3}$
6	1.0051	$5.2X10^{-3}$	$4.1X10^{-3}$
7	1.0053	$5.8X10^{-3}$	$4.3X10^{-3}$
8	1.0049	$4.9X10^{-3}$	$4.0X10^{-3}$

(b) The Face-Class Samples

Sample No.	V_1	V_2	V_3
1	$5.6X10^{-3}$	1.0058	$4.6X10^{-3}$
2	$5.5X10^{-3}$	1.0057	$4.5X10^{-3}$
3	$5.5X10^{-3}$	1.0056	$4.6X10^{-3}$
4	$5.4X10^{-3}$	1.0054	$4.4X10^{-3}$
5	$5.5X10^{-3}$	1.0056	$4.5X10^{-3}$
6	$5.1X10^{-3}$	1.0052	$4.1X10^{-3}$
7	$5.4X10^{-3}$	1.0055	$4.4X10^{-3}$
8	$5.3X10^{-3}$	1.0054	$4.3X10^{-3}$

(Table 5.1 continued)

(c) The Leaf-Class Samples

Sample No.	V_1	V_2	V_3
1	$7.7X10^{-3}$	$8.0X10^{-3}$	1.0064
2	$6.8X10^{-3}$	$6.9X10^{-3}$	1.0056
3	$7.1X10^{-3}$	$7.3X10^{-3}$	1.0059
4	$3.0X10^{-3}$	$3.3X10^{-3}$	1.0024
5	$3.4X10^{-3}$	$3.6X10^{-3}$	1.0027
6	$3.9X10^{-3}$	$4.1X10^{-3}$	1.0035
7	$3.8X10^{-3}$	$3.8X10^{-3}$	1.0030
8	$4.2X10^{-3}$	$4.5X10^{-3}$	1.0035

Table 5.2 The mapping results for the noisy sample images

(a) The Bird-Class Noisy Samples

Sample No.	V_1	V_2	V_3
1	0.9783	$2.1X10^{-2}$	$8.3X10^{-2}$
2	0.5524	$5.8X10^{-2}$	$4.1X10^{-1}$
3	1.0002	$9.9X10^{-3}$	$5.2X10^{-2}$
4	1.0169	$1.5X10^{-2}$	$3.4X10^{-3}$
5	0.9856	$-7.8X10^{-3}$	0.1086
6	0.5788	$-1.4X10^{-2}$	0.3388
7	0.9672	$6.4X10^{-2}$	0.1368
8	1.1068	$9.4X10^{-3}$	0.2144

(Table 5.2 continued)

(b) The Face-Class Noisy Samples

Sample No.	V_1	V_2	V_3
1	0.4265	0.5520	0.2533
2	0.2558	0.6309	0.4484
3	0.2828	0.6653	0.3025
4	0.5257	0.5815	-8.5×10^{-3}
5	0.3847	0.5514	0.3506
6	0.3417	0.6314	0.3265
7	0.4645	0.5495	0.3596
8	0.4016	0.5318	0.3806

(c) The Leaf-Class Noisy Samples

Sample No.	V_1	V_2	V_3
1	0.3243	1.2×10^{-2}	0.9913
2	0.4831	4.8×10^{-2}	0.8743
3	0.5529	3.6×10^{-2}	0.6474
4	0.4104	1.1×10^{-2}	0.5207
5	0.3680	2.5×10^{-2}	0.6310
6	0.3673	3.5×10^{-3}	0.6004
7	0.3106	5.6×10^{-4}	0.7336
8	0.3075	-2.0×10^{-3}	0.7864

Table 5.3 The mapping results for the noise-free object images

(B=Bird-Class, F=Face-Class, L=Leaf-Class)

No.	V_1	V_2	V_3
B-1	1.0062	$6.6X10^{-3}$	$5.2X10^{-3}$
B-2	1.0047	$4.7X10^{-3}$	$3.7X10^{-3}$
B-3	1.0030	$3.3X10^{-3}$	$2.5X10^{-3}$
B-4	1.0343	$3.5X10^{-2}$	$6.4X10^{-3}$
B-5	1.0782	$7.7X10^{-3}$	$4.7X10^{-3}$
F-1	$5.5X10^{-3}$	1.0825	$4.7X10^{-3}$
F-2	$5.6X10^{-3}$	1.0684	$4.2X10^{-3}$
F-3	$6.2X10^{-3}$	1.0248	$4.8X10^{-3}$
F-4	$6.5X10^{-3}$	1.0432	$5.1X10^{-3}$
F-5	$4.8X10^{-3}$	1.1032	$5.4X10^{-3}$
L-1	$7.7X10^{-3}$	$8.0X10^{-3}$	1.0064
L-2	$6.8X10^{-3}$	$6.9X10^{-3}$	1.0056
L-3	$7.1X10^{-3}$	$7.3X10^{-3}$	1.0059
L-4	$3.0X10^{-3}$	$3.3X10^{-3}$	1.0024
L-5	$3.4X10^{-3}$	$3.6X10^{-3}$	1.0027

Table 5.4 The mapping results for the noisy object images

(B=Bird-Class, F=Face-Class, L=Leaf-Class)

No.	V_1	V_2	V_3
B-1	0.9783	$2.1X10^{-2}$	$8.3X10^{-2}$
B-2	0.5524	$5.8X10^{-2}$	0.4134
B-3	1.0002	$9.9X10^{-3}$	$5.2X10^{-2}$
B-4	0.6237	$1.2X10^{-1}$	0.3753
B-5	0.4176	$9.3X10^{-2}$	0.5718
F-1	0.4533	0.5358	0.2479
F-2	0.3743	0.6330	0.3341
F-3	0.5756	0.5885	0.3743
F-4	0.5169	0.5821	0.0372
F-5	0.3890	0.5521	0.3267
L-1	0.3243	$1.2X10^{-2}$	0.9913
L-2	0.4831	$4.8X10^{-2}$	0.8743
L-3	0.5529	$3.6X10^{-2}$	0.6474
L-4	0.4104	$1.1X10^{-2}$	0.5207
L-5	0.3680	$2.5X10^{-2}$	0.6310

Table 5.5 The mapping results for the restored object images

(B=Bird-Class, F=Face-Class, L=Leaf-Class)

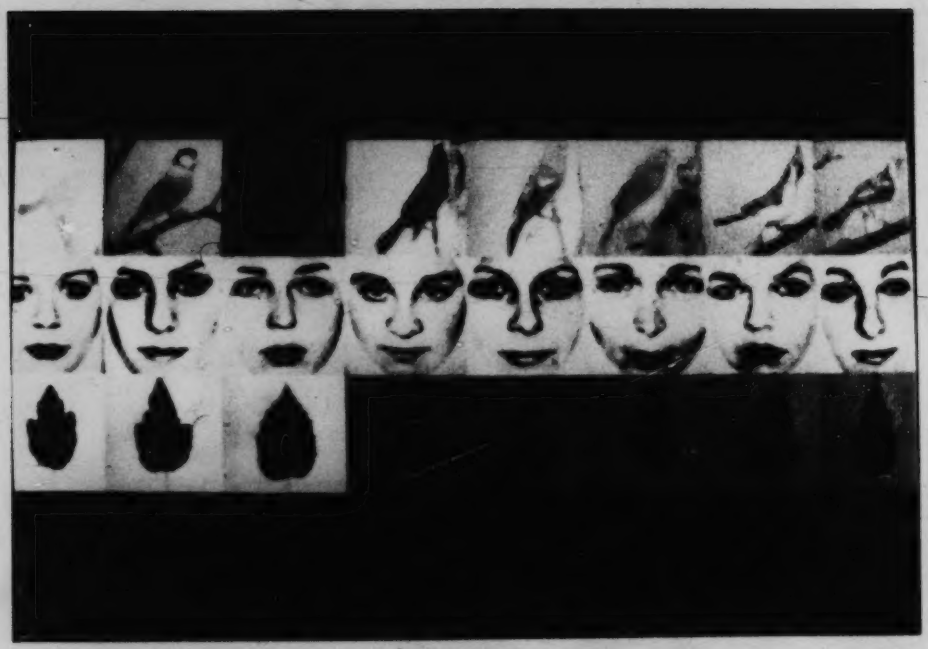
No.	V_1	V_2	V_3
B-1	0.9965	$1.8X10^{-3}$	$6.8X10^{-3}$
B-2	1.0224	$-3.4X10^{-3}$	$7.3X10^{-3}$
B-3	0.9218	$2.0X10^{-3}$	$6.5X10^{-2}$
B-4	0.9971	$8.4X10^{-3}$	$1.2X10^{-2}$
B-5	1.0874	$-1.3X10^{-3}$	$-6.8X10^{-2}$
F-1	$-6.2X10^{-3}$	1.1112	$-8.6X10^{-2}$
F-2	$1.9X10^{-2}$	1.0274	$-8.1X10^{-2}$
F-3	$-4.9X10^{-2}$	1.0518	$-5.1X10^{-3}$
F-4	$8.9X10^{-2}$	0.9390	$-1.22X10^{-2}$
F-5	$5.5X10^{-3}$	0.9828	$5.1X10^{-2}$
L-1	$-1.9X10^{-3}$	$-3.6X10^{-4}$	1.0991
L-2	$2.2X10^{-2}$	$-1.5X10^{-2}$	1.1429
L-3	$5.2X10^{-2}$	$-1.4X10^{-2}$	1.0148
L-4	$-3.1X10^{-2}4$	$2.4X10^{-2}$	0.9194
L-5	$2.3X10^{-3}$	$2.9X10^{-2}$	0.8570

Table 5.6 The cross-correlation peak values

(B=Bird-Class, F=Face-Class, L=Leaf-Class)

Object	B-1	B-2	B-3	B-4	B-5	F-1	F-2	F-3	F-4	F-5	L-1	L-2	L-3	L-4	L-5
B-1	0.92	0.63	0.54	0.57	0.62	0.21	0.20	0.18	0.16	0.22	0.23	0.20	0.22	0.24	0.21
B-2		0.89	0.55	0.63	0.67	0.16	0.20	0.20	0.13	0.19	0.22	0.21	0.20	0.24	0.26
B-3			0.93	0.54	0.73	0.18	0.19	0.21	0.18	0.17	0.12	0.15	0.14	0.18	0.18
B-4				0.93	0.65	0.20	0.18	0.19	0.16	0.18	0.16	0.17	0.20	0.19	0.23
B-5					0.95	0.18	0.18	0.20	0.17	0.16	0.17	0.16	0.14	0.14	0.16
F-1						0.92	0.73	0.73	0.70	0.72	0.23	0.23	0.21	0.22	0.20
F-2							0.90	0.72	0.73	0.73	0.21	0.21	0.20	0.18	0.20
F-3								0.89	0.72	0.70	0.21	0.23	0.19	0.19	0.22
F-4									0.93	0.74	0.19	0.23	0.21	0.20	0.18
F-5										0.92	0.23	0.21	0.18	0.20	0.19
L-1											0.88	0.75	0.80	0.74	0.77
L-2												0.92	0.76	0.81	0.78
L-3													0.90	0.80	0.78
L-4														0.92	0.76
L-5															0.90

Figure 5.7 (a) The training sample images. Three classes: BIRD class (row 1), FACE class (row 2), LEAF class (row 3). (b) The training images degraded with additive, white noise with zero mean and variance σ_v^2=1275. All the images are of 64x64 and 8-bit in resolution.



A

B

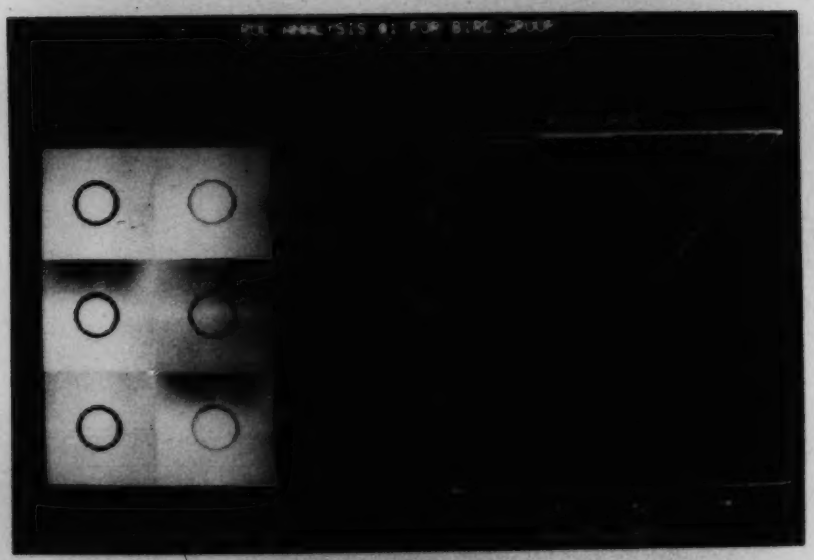
Figure 5.8 (a) 64x64 8-bit object images from the three classes. (b) The object images degraded with additive, white noise with zero mean and variance $\sigma_v^2 = 1275$.

A

B

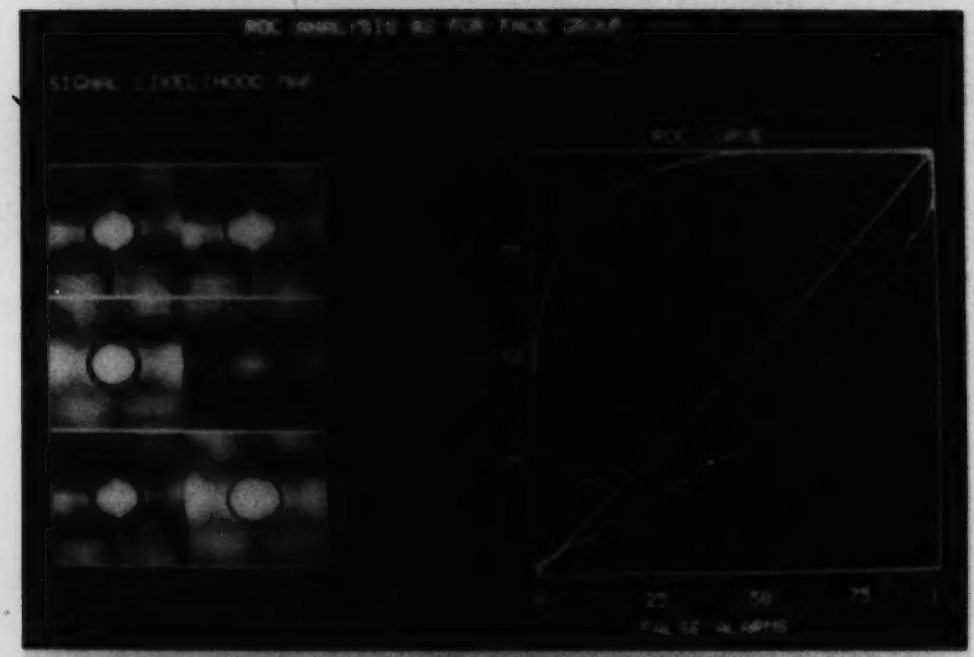
Figure 5.9 (a) The restored object images of (Figure 5.8b).
(b) The relative ROC curves for BIRD 1 (object).

A

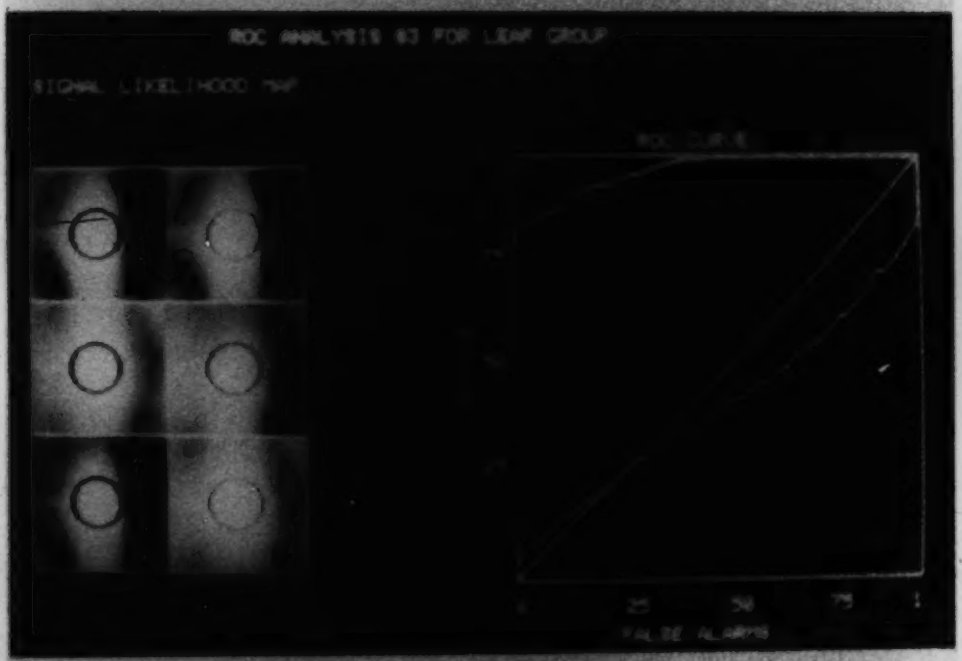


B

Figure 5.10 (a) The relative ROC curves for FACE 2 (object).
(b) The relative ROC curves for LEAF 3 (object).



A



B

CHAPTER VI

CONCLUSIONS

6.1 Summary

One of the purposes of this dissertation has been to develop some methods and computational techniques for image processing and computer vision (e.g. restoration, detection, and pattern recognition) in noisy environments. It has been shown that corruption by additive noise is one of the most common forms of degradations, and effective image processing techniques dealing with this type of noise will have significant impact on many applications. The main attention is thus restricted to images corrupted by additive noise. Several estimation techniques have been developed for restoring or detecting images corrupted by additive noise.

In Chapter II, a two-dimensional sequential Kalman filter was derived. It is well known that the Kalman filter is basically a recursive estimation technique based on a set of state and observation equations, and some à priori statistics on modeling error and observation noise. A spatial causal image field is defined which divides the entire image into three regions, namely, the "Past", the "Present", and the "Future". A linear state equation model of the original image is proposed. It has been shown in the literature (Andrews and Hunt, 1977; Woods, 1981) that the use of minimum variance as an optimality criterion on a

global setting will result in algorithms unable to respond to abrupt changes (e.g. edges) in the image. In order to handle this problem, we proposed a weighting function which varies with local statistics. It is applied as a weighting function to the Kalman filter. While the underlying idea of this approach is similar to some previously developed algorithms, the salient feature is that the Kalman filter is derived based on a stationary image model. This makes the derivation straightforward and simple. The actual nonstationarity of the image field is handled by local statistics measurement. Consequently, the algorithm is able to adjust itself to local spatial characteristics and to improve the visual quality of the restored images.

In image restoration, the restoring procedure is usually done off-line. That is, the entire image to be processed is available. This enables the use of noncausal image modeling approach. Chapter III presents an iterative adaptive restoration algorithm based on a noncausal image model. The pixel under consideration is depicted by its surrounding neighbour pixels. An observation model similar to that used in Chapter II is modified to include the local statistical information. Two optimality criteria, one for the modeling parameters, the other for the filter gain, are used. The advantage of using two separate criteria is that the parameter identification and the filter gain can be computed explicitly which makes the design of the filter more flexible. A simple local statistics measurement scheme

is employed. Because of the inclusion of the local statistical information in the observed image model, the resulting filter gain varies with the local spatial activity. The simulations show that this algorithm preserves edge information while suppresses noise in flat regions. When the gain is set to a constant the restored images are very smooth and blurred due to the smeared edges. From the comparison, it strongly suggests that images restored by edge preserving filters have better visual quality than those restored by filters developed based on global stationarity assumptions.

When an image is embedded in a nonstationary, nonwhite noise background, the detection of the image using matched filtering method will be nonoptimal. The estimation techniques introduced in Chapters II and III can be extended to detection problems. In Chapter IV, a signal reference adaptive postfiltering technique was developed to improve the detectability of the matched filter (cross-correlator). This was done by deriving an optimal filter based on the minimum variance between the signal auto-correlation and the estimated auto-correlation.

It is well known that the matched filter is simple to implement and is optimal if the background noise is stationary and white. However, in general, the optimal detector requires the full à priori knowledge of the statistics and the inverse of the noise covariance matrix.

The latter makes the computation impractical since the dimension of the covariance matrix is very large even for images of moderate sizes. It is worth noting that in derivation of the postfilter the statistical characteritics of the noise are not assumed to be known except being additive, which makes this filter more flexible in applications. Another advantage of the adaptive postfiltering technique is that the resulting algorithm does not require the inversion of the covariance matrix. The simulations show that the detectability of the matched filter is very close to that of the perfect match (i.e. the auto-correlation).

One of the common problems with computer vision is to identify objects in multi-object images which sometimes suffer from noise degradation. This problem can be and has been solved by using various detectors such as matched filters and template matching techniques (Rosenfeld and Kak, 1982). However, the matching procedures usually involve redundant computations. A number of techniques has been proposed for the detectors to get around this drawback. These, in general, follow three basic approaches:

a) The coarse-fine approach which calculates matching initially on images with reduced resolution. A full matching is performed only when a promising area is detected. In this way computation is reduced to some extent.

b) Sequential approaches in which the mismatch between

the template and the searched area is stored. As the mismatch usually grows faster than the match, the unmatched area can be rejected earlier to save computation time.

c) The edge-only technique which reduces the original image to its edge version on which matching is performed by using simple array processors. This approach significantly reduces the computation time.

In Chapter V a hierarchical approach is proposed mainly to reduce the unnecessary matching caused by the brute force application of matched filters in a multi-object detection problem. The basic philosophy behind this approach is to "divide and conquer". That is, the detection is performed in three steps, namely, (1) prefiltering to suppress noise; (2) classification by using pattern recognition method to classify the objects; (3) final detection by applying the matched filtering technique only to those objects belonging to the same class as the signal. Since in the context of signal detection the visual quality of the image is not a main concern so long as the false alarm rate is kept low, a very simple and fast averging prefilter is derived. The simulations demonstrate the effectiveness of this approach.

During last two decades many image processing techniques have been developed. The Kalman filtering technique is also employed in image restoration. For the most part, however, the two-dimensional Kalman filtering algorithms have treated images as stationary image fields.

Only recently, a few Kalman filtering techniques have been proposed to deal with the nonstationarity of images. The basic principles involved in this approach are as follows,

a) Partition the image into subregions each of which has stationary statistics (e.g. constant variance). A bank of Kalman filters is designed for the subregions. To improve visual quality, a human visual model is incorporated into the Kalman filter.

b) Assume nonstationary statistics (nonstationary mean, or nonstationary variance, or both) for the image field, and design a global Kalman filter based on this assumption.

In this thesis the two-dimensional Kalman filter is first derived based on the stationarity assumption. The local statistics are then used to modify the resulting Kalman filter to handle the nonstationarity problem. This is a distinctly different approach and as shown in the simulations, this Kalman filter is very effective. The implementation of such a Kalman filter is also rather simple. Another difference in the proposed iterative algorithm is that the local statistics which represent the nonstationarity of the image are included in the output model rather than, as many other researchers have done, in the assumption or the image model.

It should be pointed out that although the idea of signal reference adaptive postfiltering technique originated from the control system identification theory, the

postfiltering approach proposed in Chapter IV for improving two-dimensional detection of object in nonwhite, nonstationary noisy environments is a new concept. This approach is simple to use and the simulations show that the results are quite satisfactory.

6.2 Recommendations for Further Research

This thesis has been devoted entirely to the development of algorithms to process images degraded by additive noise. As mentioned in Chapter I, images also suffer from other forms of degradations. For example, out-of-focus will result in blurred images, and some noise is signal dependent. Thus image processing techniques developed in this thesis will be potentially useful if these algorithms are improved or modified to take into account the difficulties metioned above.

An immediate modification which comes to mind is to consider the point spread function as a linear polynomial with unknown parameters. This problem has been considered by others, however, for the most part, the point spread functions are assumed to be of some very trivial forms. The difficulty with this is that the restoration is essentially an inverse problem which requires inferring the input from the output information. Thus it is more realistic to derive some PSF's (point spread functions) by using experiments in different applications. With explicit forms of PSF's, the existing restoration algorithms can be improved to handle

the blurring effect as well as additive noise. Other modifications would be to include the signal dependent noise in the process. These modifications will inevitably increase the complexity of the algorithms, and are likely to be application dependent.

Developing an adaptive postfilter for the matched filter to detect images degraded by blurring, and signal dependent noise in addition to additive noise could be quite useful. As discussed in Chapter IV, in image detection both the object and the output are known, so that certain techniques in control system identification theory could be extended to estimate the parameters of the imaging systems (e.g. PSF's) and a postfilter could then be designed to improve the matched filter in a changing environment.

It has been demonstrated (Caelli and Nagendran, 1986) that using edge-only matching technique to detect an object is computationally realistic and efficient. This method, however, has a major drawback—ambiguity. Since a very large number of different objects have the same edge profile, this procedure will often lead to false alarms. This problem may be alleviated to some extent by using some of the extracted image features in addition to the edge-only information. The feature extraction procedure should, however, be simple if the overall efficiency of the detection process is to be preserved. One approach which immediately comes to mind is to regard the image field as a matrix and use its singular

values as the feature which at the most is a set of 512 singular values (for a 512x512 image). The edge-only matching need be performed only when two images have singular values similar to each other. This procedure will, hopefully, reduce the probability of false alarms. It is also reasonable to expect that the hierarchical multi-object detection approach can be applied to edge-only images.

One of the problems with pattern recognition is that many existing techniques are rotation variant. In pattern recognition literature, methods are developed under the assumption (sometimes, implicitly) that the objects from a given class possess the same orientation as the training samples. Most of these methods perform poorly when the orientation of the objects are different from that of the training samples. Although, in some applications, such an orientation dependency property is desirable, e.g., discriminating between similar textures which are oriented differently, in general, it reduces the flexibility of the method. It is a very interesting topic to develop some pattern recognition algorithms with rotation invariant property. That is, the results of pattern recognition is not affected by the rotation of the objects. Although reseach in this direction has been carried out recently, the techniques are mostly for texture classification and the accuracy of these methods is generally not high.

Three dimensional image processing and vision computing

have gained more attention recently, since three dimensional images are the most realistic visual forms in many applications, particularly, in computer vision. The restoration algorithms developed in this thesis are applicable to three dimensional images. It is an extremely promising research area where new approaches are needed, since direct extension of the existing techniques in two-dimensional image processing to three dimensions is not likely to work properly in general. It is the complicated and varied geometry of three-dimensional images that makes the three dimensional image processing more challenging.

REFERENCES

Aboutalib A. O. et al. (1972). Digital restoration of images degraded by general motion blur. IEEE Trans. Automat. Contr., AC-2, No. 3, 294-302.

Abramatic J. F. and L. M. Silverman (1982). Nonlinear restoration of noisy images. IEEE Trans. Pattern Recog. Mach. Intell., PAMI-4, No. 2, 141-149.

Abu-Mostafa Y. S. and D. Psaltis (1985). Image normalization by complex moments. IEEE Trans. Patt. Analy. Machine Intellig., PAMI-7, No. 1, 46-55.

Ahmed N. and K. R. Rao (1975). Orthogonal transforms for digital signal processing. Springer-Verlag, Berlin.

Anderson B. D. O. and J. B. Moore (1972). The Kalman-Bucy filter as a true time-varying Wiener filter. IEEE Trans. Syst., Man, Cybern., SMC-1, No. 2, 119-128.

Anderson B. D. O. and J. B. Moore (1979). Optimal filtering. Prentice-Hall, Englewood Cliffs, N.J..

Anderson G. L. and A. N. Netravali (1976). Image restoration based on a subjective criterion. IEEE Trans. Syst., Man, Cybern., SMC-6, No. 6, 845-853.

Andrews H. C.(1974). Digital image processing: A Survey. Computer, Vol. 7, No. 5, 34-45.

Andrews H. C. and B. R. Hunt (1977). Digital image restoration. Prentice-Hall, Englewood Cliffs, New Jersey.

Arcese A. et al. (1970). Image detection through bipolar correlation. IEEE Trans. Inform. Theory, IT-16, No. 5, 534-541.

Barb D. F. (1975). Image devices using the charge-coupled concept. Proc. IEEE, Vol. 63, 38-67.

Biemond J., J. Rieske, and J. J. Gerbrands (1983). A Fast Kalman filter for image degraded by both blur and noise. IEEE Trans. Acoust., Speech, Signal Processing, ASSP-31, No. 5, 1248-1255.

Budrikis Z. L. (1972). Visual fidelity criterion and modeling. Proc. IEEE, Vol. 60, 771-779.

Caelli T. M. and G. Moraglia, On the detection of signals embedded in natural scenes. Perception and Psychophsics,

Vol. 39, No. 2, 87-95, (1986).

Caelli T. M. and S. Nagendran, Fast edge-only matching techniques for robot pattern recognition. Comput. Graph. and Image Processing, in press, (1986).

Chan P. and J. S. Lim (1984). One-dimensional processing for adaptive image restoration. Proc. IEEE, Int. Conf. Acoust., Speech, Signal Processing, ICASSP-84, San Diego, CA., 37.3.1-37.3.4.

Chien Y. T. and K. S. Fu (1967). On the generalized Karhunen-Loeve expansion. IEEE Trans. Inform. Theory, IT-13, No. 4, 518-520.

Cormack A. M. (1963). Representation of a function by its line integrals, with some radiological applications. J. Appl. Phys., Vol. 34, No. 9, 2722-2727.

Dessimoz J.-D. et al. (1984). Matched filters for bin picking. IEEE Trans. Patt. Analy. Machine Intellig., PAMI-6, No. 6, 686-697.

Devijer P. A. and J. Kittler (1982). Pattern recognition: A Statistical approach. Prentice-Hall International, London.

Duda R. O. and P. E. Hart (1973). Pattern classification and scene analysis. Wiley-Interscience, John Wiley & Sons, New York.

Dudani S. et al. (1977). Aircraft identification by moment invariants. IEEE Trans. Comput., C-26, 39-45.

Dudgeon D. E. (1980). An Iterative implementation of 2-D digital filters. IEEE Trans. Acoust., Speech, Signal Processing, ASSP-28, No. 6, 666-671.

Ekstrom M. P. and J. W. Woods (1976). Two-dimensional spectral factorization with applications in recursive digital filtering. IEEE Trans. Acoust., Speech, Signal Processing, ASSP-24, No. 2, 115-128.

Foley D. H. and J. W. Sammom, Jr. (1975). An Optimal set of discriminant vectors. IEEE Trans. Comput., C-24, No. 3, 281-289.

Franks L. E. (1966). A Model for the random video process. Bell Syst. Tech. J., Vol. 45, 609-630.

Fu K. S. (1982). Syntactic pattern recognition and applications. Prentice-Hall Inc., Englewood Cliffs, (New Jersey, 1982).

Fukunaga K. and W. L. G. Koontz (1970). Application of the Karhunen-Loeve expansion to feature selecton and ordering. IEEE Trans. Comput., C-19, No. 4, 311-318.

Gonzalez R. and P. Wintz (1977). Digital image processing. Addison-Wesley, Reading, MA..

Gray R. M. (1971). Toeplitz and circulant matrices: A review. Tech. Rep. No.6502-1, Center for Systems Research, Stanford Univ.

Green W. B. (1983). Digital image processing: A systems approach. Van Nostrand Reinhold Company, New York.

Habibi A. (1972). Two-dimensional Bayesian estimation of images. Proc. IEEE, Vol. 60, 878-883.

Hsu Y.-N. et al. (1982). Rotation-invariant digital pattern recognition using circular harmonic expansion. Appl. Opt., Vol. 21, No. 22, 4012-4015.

Hsu Y.-N. et al. (1982a). Optical pattern recognition using circular harmonic expansion. Appl. Opt., Vol. 21, No. 22, 4016-4019.

Hsu Y.-N. et al. (1984). Pattern recognition by multiple circular harmonic expansion. Appl. Opt., Vol. 23, No. 6, 841-844.

Hu M.-K. (1962). Visual pattern recognition by moment invariants. IRE Trans. Inform. Theory, IT-8, 179-187.

Huang T.S. (1966). Some notes on film-grain noise. NSF Summer Study Rep., Woods Hole, MA., 105-109.

Hunt B. R. (1977). Bayesian methods in nonlinear digital image restoration. IEEE Trans. Comput., C-26, No. 3, 219-229.

Hunt B. R. and T. M. Cannon (1976). Nonstationary assumptions for Gaussian models of images. IEEE Trans. Syst., Man, Cybern., SMC-6, No. 12, 876-881.

Ingle V. K., A. Radpour, J. W. Woods, and H. Kaufman (1978). Recursive estimation with non-homogeneous image models. Proc. IEEE Pattern Recog. Image Processing, Chicago, IL., 105-108.

Ingle V. K. and J. W. Woods (1976). Multiple image recursive estimation images. Proc. IEEE, Int. Conf. Acoust., Speech, Signal Processing, ICASSP-78, Washington, D.C., 642-645.

Kalman R. E. (1960). A new approach to linear filtering and

prediction problems. J. Basic Engineering, (March, ASME Trans.), 35-45.

Kasturi R. et al. (1985). Adaptive point estimation in signal-dependent noise. IEEE Trans. Syst. Man, Cybern., SMC-15, No. 3, 352-361.

Kaufman H., J. W. Woods, S. Dravida, and A. M. Tekalp (1983). Estimation and identification of two-dimensional images. IEEE Trans. Automat. Contr., AC-28, No. 7, 745-756.

Kondo H. et al. (1981). Improved method of Bayesian estimation for nonlinear image restoration. Trans. Inst. Electron. Commun., Eng. Jnp. Sect. E(Engl.), V. E64, No. 11, 744-745.

Kuan D. T., A. A. Sawchuk, T. C. Strand, and P. Chavel (1984). Nonstationary 2-D recursive restoration of images with signal-dependent noise. Proc. Int. Conf. Acoust., Speech, Signal Processing, ICASSP-84, San Diego, CA., 37.4.1-37.4.4.

Lanczos C. (1961). Linear differential operators. D. Van Nostrand Comp. LTD., London.

Lebedev D. S. and L. I. Mirkin (1975). Smoothing of two-dimensional images using the 'composite' model of a fragment. Iconics-Digital Holography-Image processing, USSR Inst. Inform. Transmission Problems, Academic of Sciences, 57-62.

Lee J.-S. (1980). Digital image enhancement and noise filtering by use of local statistics. IEEE Trans. Patt. Analy. Machine Intellig., PAMI-2, No. 2, 165-168.

Lee J.-S. (1981). Refined filtering of image noise using local statistics. Computer Graph. and Image processing, Vol. 15, 380-389.

Limb J. O. and C. B. Rubenstein (1978). On the Design of quantizers for DPCM coders: a functional relationship between visibility, probability and masking. IEEE Trans. Commun., COM-26, No. 2, 573-578.

Liu Z.-Q. and T. M. Caelli (1986). An Adaptive postfiltering technique for detecting image objects in non-white noise. Applied Optics, May 15, 1986, Vol. 25, No. 10, pp. 1622-1626.

Ljung L. and T. Söderström (1983). Theory and practice of recursive identification. MIT press, Cambridge, MA..

Lo C. M. and A. A. Sawchuk (1979). Nonlinear restoration of

filtered images with Poisson noise. Proc. IEEE Tech. Symp. Appl. Digital Image Processing-III, San Diego, CA., Vol. 207, 84-95.

Maitra S. (1979). Moment invariants. Proc. IEEE, Vol. 67, No. 4, 697-699.

Munteanu C. (1981). Evaluation of the sequential similarity detection algorithm applied to binary images. Pattern Recognition, Vol. 13, No. 2,167-175.

Murphy M. S. and L. M. Silverman (1978). Image model representation and line-by-line recursive restoration. IEEE Trans. Automat. Contr., AC-23, No. 5, 809-816.

Naderi F. and A. A. Sawchuk (1978). Estimation of images degraded by film-grain noise. Appl. Opt., Vol. 17, 1228-1237.

Nahi N. E. (1972). Role of the recursive estimation in statistical image enhancement. Proc. IEEE, Vol. 60, No. 7, 872-877.

Nahi N. E. and A. Habibi (1975). Decision-directed image enhancement. IEEE Trans. Circui. Syst., CAS-22, No. 3, 286-293.

Netravali A. N. and B. Brasada (1975). Adaptive quantization of picture signals based on spatial masking. Proc. IEEE, Vol. 65, No. 4, 536-548.

Niemann H. (1979). Digital image analysis. in Advances in Digital Image Processing, Ed. by P. Stucki, 77-122.

Papoulis A. (1984). Probability, random variables, and stochastic processes. 2nd edition, McGraw-Hill Book Company, New York.

Powell S. R. and L. M. Silverman (1974). Modeling of two-dimensional covariance functions with application to image restoration. IEEE Trans. Automat. Contr., AC-19, No. 1, 8-13.

Pratt W. K. (1978). Digital image processing. John Wiley and Sons, New York.

Rajala S. A. and R. P. DeFigueiredo (1981). Adaptive nonlinear image restoration by a modified Kalman filtering approach. IEEE Trans. Acoust., Speech, Signal Processing, ASSP-29, No. 5, 1033-1042.

Rino C. L. (1970). The Inversion of matrices by finite Fourier transforms. IEEE Trans. Inform. Theory, IT-16, No. 2, 230-232.

Rosenfeld A. and A. C. Kak (1982). Digital picture processing. Academic press, 2nd Edition, New York.

Rosenfeld A. and G. J. Vanderdrug (1977). Coarse-fine template matching. IEEE Trans. Syst., Man, Cybern., SMC-7, No. 2, 104-107.

Sage A. P. and J. L. Melsa (1971). Estimation theory. McGraw-Hill Book Company, New York.

Sammon, Jr. J. W. (1970). Interactive pattern analysis and classification. IEEE Trans. Comput., C-19, No. 7, 594-616.

Schafer R. W. and R. M. Mersereau (1981). Constrained iterative restoration algorithms. Proc. IEEE, Vol. 69, 432-450.

Tanimoto S. L. (1981). Template Matching in pyramids. Comput. Graph. and Image Processing, Vol.16, 356-369.

Teague M. (1980). Image analysis via the general theory of moments. J. Opt. Soc. Am. (A), Vol. 77, 920-930.

Trussell H. J. and B. R. Hunt (1978). Sectioned methods for image restoration. IEEE Trans. Acoust., Speech, Signal Processing, ASSP-26, No. 2, 157-164.

Trussell H. J. and B. R. Hunt (1979). Improved methods of maximum à postériori image restoration. IEEE Trans. Comput., C-28, No. 1, 57-62.

Turin G. L. (1976). An Introduction to digital matched filters. Proc. IEEE, Vol. 64, No. 7, 1092-1112.

Vanderbrug G. J. and A. Rosenfeld (1977). Two-stage template matching. IEEE Trans. Comput., C-26, No. 4, 384-393.

Van Trees H. L. (1968). Detection, estimation, and modulation theory. Part 1. John Wiley & Sons, Inc., New york.

Vijaya Kumar B. V. K. and C. Carroll (1984). Loss of optimality in cross-correlators. J. Opt. Soc. Am. (A), Vol. 1, No. 4, 392-397.

Wallis R. (1976). An Approach to the space variant restoration and enhancement of images. Proc. Symp. Current Math. Problems in Image Science, Naval Postgraduate School, Monterery, CA., Ed. C. O. Wilde and E. Barrett, 107-111.

Ward R. K. and B. E. A. Saleh (1985). Restoration of images

distorted by systems of random impulse response. J. Opt. Soc. Am. (A), Vol. 2, No. 8, 1254-1259.

Watanabe S. (1965). Karhunen-Loeve expansion and factor analysis. Trans. 4th Prague Conf. Inform. Theory.

Willsky A. S. (1978). Relationships between digital signal processing and control and estimation theory. IEEE Proc. Vol. 66, No. 9, 996-1017.

Wong R. Y. and E. L. Hall (1978). Sequential hierarchical scene matching. IEEE Trans. Comput., C-27, 359-366

Woods J. W. (1979). Correction to 'Kalman filtering in two dimensions' IEEE Trans. Infor. Theory, IT-25, No. 5, 628-629.

Woods J. W. (1981). Two-dimensional Kalman filtering. in Two-Dimensional Digital Signal Processing: I, S. T. Huang ed., Springer-Verlag, New York, 155-205.

Woods J. W. and V. K. Ingle (1981a). Kalman filtering in two dimensions: Further results. IEEE Trans. Acoust., Speech, Signal Processing, ASSP-29, No. 2, 188-197.

Woods J. W. and C. H. Radewan (1976). Reduced update Kalman filter: A two-dimensional recursive processor. John Hopkins Conf. Inform. Sci. Syst., Baltimore, MD..

Woods J. W. and C. H. Radewan (1977). Kalman filtering in two dimensions. IEEE Trans. Infor. Theory, IT-23, No. 4, 473-482.

Yamakoshi Y. and T. Sato (1982). Iterative image restoration from data available in multiple restricted regions. Appl. Opt., Vol. 21, No. 24, 4473-4480.

APPENDIX I

Random Fields

1. Definition of a Random Field

We consider a family of functions $f(p,\omega_i)$ over the set of all outcomes $\Omega=\{\omega_1,\omega_2,\ldots\}$ of an experiment, where $p \in P$, and P is an interval on the real axis or a region of a multidimensional Euclidean space. When P is one-dimansional, the family of functions $f(p,\omega_i)$ is called a stochastic process. When the dimensionality of P is greater than one, the family of functions $f(p,\omega_i)$ is defined as a *random field*.

In image processing, images are considered as two-dimensional random fields. In the sequel, we shall discuss some properties of random fields with P being two-dimensional.

Let P define the XY-plane, so that p is a point in the XY-plane, and p can be represented by its coordinates, i.e., $p=(x,y)$. With this notation, the random field $f(p,\omega_i)$ can be expressed as $f[(x,y),\omega_i]$. For a given value of (x,y), $f[(x,y),\omega_i]$ is a random variable, while for a given outcome ω_i, $f[(x,y),\omega_i]$ is a function over the XY-plane. This is illustrated in Figure A1.1, with $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$. The underlying experiment could, for example, be samples of images from a collection. The outcome ω_i corresponds to the

sample of the i-th image.

Obviously, $f[(x,y),\omega_i]$ is a function of ω_i (see Figure A1.1), therefore, a random variable for a fixed (x,y). On the other hand, for a fixed ω_i, $f[(x,y),\omega_i]$ is a two-dimensional function (or, image) in the XY-plane.

The random field $f[(x,y),\omega_i]$ is a random variable for a specific value of (x,y). In general, this random variable will not have the same statistical properties for all values of (x,y). In other words, the distribution and density functions for the random variable $f[(x,y),\omega_i]$ will depend on the value of (x,y).

As with one-dimensional random variables, we define two-dimensional distribution function at (x,y) in the XY-plane as

$$P_f[t;(x,y)] \triangleq p\{f[(x,y),\omega_i] \leq t\}. \qquad \text{(A1.1)}$$

The density function at (x,y) is given by

$$p_f[t;(x,y)] \triangleq \partial P_f[t;(x,y)]/\partial t. \qquad \text{(A1.2)}$$

For notational simplicity, we shall denote $f[(x,y),\omega_i]$ by $f(x,y)$.

Given points (x_1,y_1), (x_2,y_2), ..., (x_N,y_N). The joint distribution and density functions of these random variables

are given by

$$P_t[t_1,t_2,\ldots,t_N;(x_1,y_1),(x_2,y_2),\ldots,(x_N,y_N)]$$
$$\triangleq P\{f(x_1,y_1)\leq t_1, f(x_2,y_2)\leq t_2,\ldots,f(x_N,y_N)\leq t_N\}, \qquad \text{(A1.3)}$$

and

$$p_t\,[t_1,t_2,\ldots,t_N;(x_1,y_1),(x_2,y_2),\ldots,(x_N,y_N)]$$
$$\triangleq \frac{\partial P_t[t_1,t_2,\ldots,t_N;(x_1,y_1),(x_2,y_2),\ldots,(x_N,y_N)]}{\partial t_1,\partial t_2,\ldots,\partial t_N}, \qquad \text{(A1.4)}$$

respectively.

2. Mean, Correlation, and Covariance

It is known that, in general, the density function of a random variable depends on the value of (x,y) chosen. As a consequence, its expected value must also be a function of (x,y). Let $\mu_t(x,y)$ denote this expectation, thus

$$\mu_t(x,y) = E\{f(x,y)\} = \int_{-\infty}^{\infty} t p_t[t;(x,y)]dt. \qquad \text{(A1.5)}$$

The correlation of two random fields f(x,y) and g(x,y) is defined as the expected value of the product of the two

random variables $f(x_1,y_1)$ and $g(x_2,y_2)$:

$$R_{fg}(x_1,y_1;x_2,y_2) = E\{f(x_1,y_1)g(x_2,y_2)\}$$
$$= \int_{-\infty}^{\infty}\int t_1 t_2 p(t_1,t_2;x_1,y_1,x_2,y_2)dt_1dt_2. \qquad (A1.6)$$

The covariance of two random fields is defined as

$$Cov_{fg}(x_1,y_1;x_2,y_2) = E\{[f(x_1,y_1)-\mu_f(x_1,y_1)][g(x_2,y_2)-\mu_g(x_2,y_2)]\}$$
$$= R_{fg}(x_1,y_1;x_2,y_2)-\mu_f(x_1,y_1)\mu_g(x_2,y_2). \qquad (A1.7)$$

If $Cov_{fg}(x_1,y_1;x_2,y_2)=0$ for all (x_1,y_1) and (x_2,y_2), the two random fields $f(x_1,y_1)$ and $g(x_2,y_2)$ are said to be uncorrelated. This is, from Eqs.(A1.6) and (A1.7), equivalent to

$$E\{f(x_1,y_1)g(x_2,y_2)\} = E\{f(x_1,y_1)\}E\{g(x_2,y_2)\}. \qquad (A1.8)$$

3. Wide Sense Stationary Random Fields

The concept of two-dimensional stationary random fields is an extension of one-dimensional wide sense stationary stochastic process.

A random field is called stationary if its expected

value $\mu(x,y)$ is independent of position (x,y) in the XY-plane, i.e.,

$$\mu(x,y) = \mu, \tag{A1.9}$$

and if its autocorrelation function is shift invariant,

$$\begin{aligned} R_{ff}(x_1,y_1;x_2,y_2) &= E\{f(x_1+x_0,y_1+y_0)f(x_2+x_0,y_2+y_0)\} \\ &= R_{ff}(x_1+x_0,y_1+y_0;x_2+x_0,y_2+y_0), \end{aligned} \tag{A1.10}$$

for all (x,y) and (x_0,y_0) in the XY-plane. By setting $x_0=-x_2$, $y_0=-y_2$, and $\alpha=x_1-x_2$, $\beta=y_1-y_2$, Eq.(A1.10) can be rearranged as

$$R_{ff}(x_1,y_1;x_2,y_2) = R_{ff}(x_1-x_2,y_1-y_2) = R_{ff}(\alpha,\beta). \tag{A1.11}$$

Similarly, two random fields $f(x,y)$ and $g(x,y)$ are jointly stationary, if

$$R_{fg}(x_1,y_1;x_2,y_2) = R_{fg}(\alpha,\beta). \tag{A1.12}$$

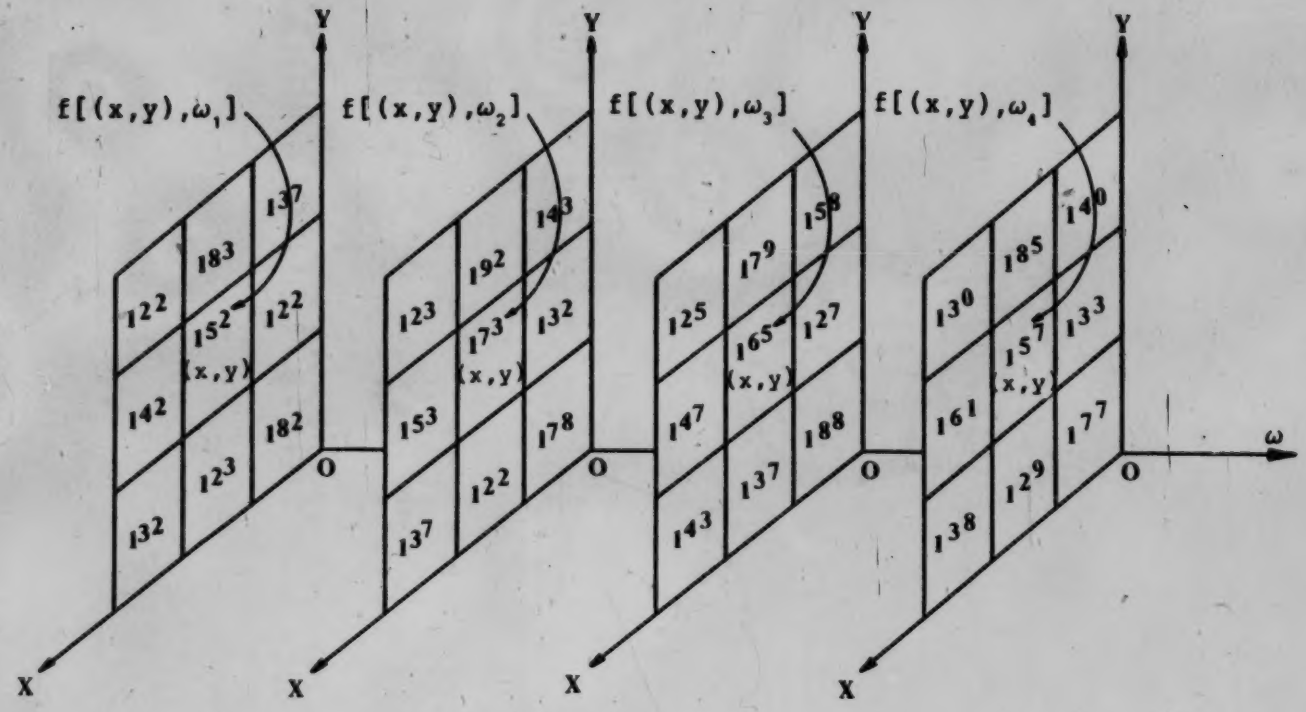


Figure A1.1 An example of 2-D random fields

APPENDIX II

The normalized cross-correlation function in Chapter IV is determined via

$$C_{rgn}(x,y) = \frac{C_{rg}(x,y)}{\sqrt{q^2(x,y)}}, \qquad (A2.1)$$

where

$$C_{rg}(x,y) = \iint_A r(u,v)g(x+u,y+v)dudv, \qquad (A2.2)$$

and

$$q^2(x,y) = \iint_A g^2(x+u,y+v)dudv, \qquad (A2.3)$$

where A denotes the integration region. These functions can be computed, using Fourier methods, by

$$C_{rg}(x,y) = F^{-1}\{R^*G\},\ R = F(r),\ G=F(g), \qquad (A2.4)$$

$$q^2(x,y) = F^{-1}\{I^*Q^2\},\ I = F(i), Q^2 = F(q^2), \qquad (A2.5)$$

where F and F^{-1} correspond to the Fourier transform and its inverse respectively. i corresponds to the unit square of

identical size as the signal. * denotes the complex conjugate.

APPENDIX III

Calculation of the Inverse Matrix Using Singular Value Decomposition (SVD)

For conciseness we define

$$T \triangleq \sum_{k=1}^{K}\sum_{l=1}^{L} f_{kl}^{(s)} f_{kl}^{(s)T} = QQ^T, \tag{A3.1}$$

where $T \epsilon R^{M^2 XN^2}$, $f_{ij}^{(s)} \epsilon R^{N^2}$, and $Q \epsilon R^{N^2 XKL}$ The matrix Q is defined by

$$Q = \left[f_{11}^{(s)}, f_{12}^{(s)}, \ldots, f_{1L}^{(s)}, f_{21}^{(s)}, f_{22}^{(s)}, \ldots, f_{2L}^{(s)}, \ldots; f_{K1}^{(s)}, f_{K2}^{(s)}, \ldots, f_{KL}^{(s)}\right]. \tag{A3.2}$$

Let

$$W = Q^T Q, \tag{A3.3}$$

we know that $W \epsilon R^{KLXKL}$ and in general, $KL << N^2$. Using the singular value decomposition technique (Lanczos, 1961) we can write (A3.3) as follows

$$W = \Phi \Sigma \Phi^T, \tag{A3.4}$$

where Φ is an orthonormal matrix i.e. $\Phi^T\Phi = \Phi\Phi^T = I$, and Σ is a KLXKL diagnal matrix consisting of the eigenvalues of W.

From (A3.4) we have

$$W\Phi = \Phi\Sigma \tag{A3.5}$$

which, from (A3.3), is

$$Q^TQ\Phi = \Phi\Sigma. \tag{A3.6}$$

Premultiplying both sides of (A3.6) by Q to yields

$$(QQ^T)Q\Phi = Q\Phi\Sigma. \tag{A3.7}$$

We know that QQ^T is an N^2XN^2 real symmetric matrix for which there exists an orthogonal matrix, such that QQ^T is diagonal. From (A3.7) we find that $Q\Phi$ is actually such an orthogonal matrix which contains KL eigenvectors of QQ^T. For our purpose, it is necessary to find the orthonormal matrix. Since

$$(QQ^T)Q\Phi = \Phi^TQ^TQ\Phi, \tag{A3.8}$$

from (A3.3) and Φ is an orthonormal matrix, we obtain

$$(Q\Phi)^TQ\Phi = \Sigma. \tag{A3.9}$$

This can be easily normalized as follows

$$\Sigma^{-1/2}[(Q\Phi)^TQ\Phi]\Sigma^{-1/2} = I. \tag{A3.10}$$

Eq.(A3.10) can be rearranged as

$$(Q\Phi\Sigma^{-1/2})^T(Q\Phi\Sigma^{-1/2}) = I. \tag{A3.11}$$

Define

$$U = Q\Phi\Sigma^{-1/2}, \tag{A3.12}$$

which is the desired orthonormal matrix. In view of Eq.(A3.11), clearly Eq.(A3.7) is equivalent to

$$(QQ^T)Q\Phi\Sigma^{-1/2} = (Q\Phi\Sigma^{-1/2})\Sigma, \tag{A3.13}$$

or

$$QQ^T = U\Sigma U^T. \tag{A3.14}$$

From (A3.1) we have

$$T = U\Sigma U^T, \tag{A3.15}$$

Thus the inverse matrix T^{-1} is given by

$$T^{-1} = U\Sigma^{-1}U^T = Q\Phi\Sigma^{-2}\Phi^TQ^T. \tag{A3.16}$$